A cross-modal, cross-species comparison of connectivity measures in the primate brain

Andrew T. Reid[a,*], John Lewis[b], Gleb Bezgin[c], Budhachandra Khundrakpam[b], Simon B. Eickhoff[a,e], Anthony R. McIntosh[c], Pierre Bellec[d], Alan C. Evans[b]

[a] *Institute of Neuroscience and Medicine (INM-1), Research Centre Jülich, Jülich, Germany*
[b] *McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, Canada*
[c] *Rotman Research Institute of Baycrest Centre, University of Toronto, Toronto, ON, Canada*
[d] *Centre de recherche de l'institut de Gériatrie de Montréal CRIUGM, Montreal, QC, Canada*
[e] *Institute of Clinical Neuroscience and Medical Psychology, Heinrich Heine University, Düsseldorf, Germany*

Abstract

In systems neuroscience, the term "connectivity" has been defined in numerous ways, according to the particular empirical modality from which it is derived. Due to large differences in the phenomena measured by these modalities, the assumptions necessary to make inferences about axonal connections, and the limitations accompanying each, brain connectivity remains an elusive concept. Despite this, only a handful of studies have directly compared connectivity as inferred from multiple modalities, and there remains much ambiguity over what the term is actually referring to as a biological construct. Here, we perform a direct comparison based on the high-resolution and high-contrast Enhanced Nathan Klein Institute (NKI) Rockland Sample neuroimaging dataset, and the CoCoMac database of tract tracing studies. We compare four types of commonly-used primate connectivity analyses: tract tracing experiments, compiled in CoCoMac; group-wise correlation of cortical thickness; tractographic networks computed from diffusion-weighted MRI (DWI); and correlational networks obtained from resting-state BOLD (fMRI). We find generally poor correspondence between all four modalities, in terms of correlated edge weights, binarized comparisons of thresholded networks, and clustering patterns. fMRI and DWI had the best agreement, followed by DWI and CoCoMac, while other comparisons showed striking divergence. Networks had the best correspondence for local ipsilateral and homotopic contralateral connections, and the worst correspondence for long-range and heterotopic contralateral connections. K-means clustering highlighted the lowest cross-modal and cross-species consensus in lateral and medial temporal lobe, anterior cingulate, and the temporoparietal junction. Comparing the NKI results to those of the lower resolution/contrast International Consortium for Brain Imaging (ICBM) dataset, we find that the relative pattern of intermodal relationships is preserved, but the correspondence between human imaging connectomes is substantially better for NKI. These findings caution against using "connectivity" as an umbrella term for results derived from single empirical modalities, and suggest that any interpretation of these results should account for (and ideally help explain) the lack of multimodal correspondence.

Keywords: connectomics, tract tracing, CoCoMac, resting state functional MRI, cortical thickness, structural covariance, diffusion-weighted MRI, hierarchical clustering, k-means clustering

*Corresponding author
Email addresses: a.reid@fz-juelich.de (Andrew T. Reid), jlewis@bic.mni.mcgill.ca (John Lewis), gbezgin@rotman-baycrest.on.ca (Gleb Bezgin), budha@bic.mni.mcgill.ca (Budhachandra Khundrakpam), s.eickhoff@fz-juelich.de (Simon B. Eickhoff), rmcintosh@rotman-baycrest.on.ca (Anthony R. McIntosh), pierre.bellec@criugm.qc.ca (Pierre Bellec), alan@bic.mni.mcgill.ca (Alan C. Evans)

1. Introduction

Characterizing the connectivity of the human brain has been a pursuit of neuroscientists since the basic concepts of axon and synapse were proposed over a century ago. Recent technological advances in both microcircuitry analysis and noninvasive imaging have given new impetus to the study of connectivity. This emerging discipline, now popularly referred to as "connectomics", ranges from single synapses, to local microcircuitry, to whole-brain networks comprised of anatomically distinct brain regions (reviewed in Behrens and Sporns, 2012). In the latter case, studies have been traditionally confined to tract tracing experiments in nonhuman primates, in which a histological tracer substance is injected in vivo and transported along axonal fibres in either anterograde or retrograde directions, labelling the set of regions that the injected site projects to, or receives projections from, respectively. The literature covering such experiments is expansive, but has been systematically collated into a single database called CoCoMac (Stephan et al., 2001; Kötter, 2004; Bakker et al., 2012). While histological tract tracing remains the gold standard for connectivity in the monkey brain, its applicability to human anatomy relies upon an assumption of homology which becomes increasingly problematic for regions with more phylogenetic divergence, such as the frontal and parietal lobes (Zilles et al., 1988; Rilling and Seligman, 2002; Rilling et al., 2008; Petrides et al., 2012). As a consequence, direct evidence for human connectivity remains elusive.

The advent of high-resolution neuroimaging techniques such as magnetic resonance imaging (MRI), however, has provided researchers with a variety of noninvasive tools with which to investigate the human connectome in an indirect way. Diffusion-weighted imaging (DWI) detects the isotropy of molecular diffusion in brain tissue, which carries information about the orientation and microstructure of white matter (WM) fibres. This approach is the basis of *probabilistic tractography*, in which multiple tracts originating in a specific seed point are generated by sampling the distribution of orientations within each subsequent voxel, until a predefined target mask is encountered. Tract counts obtained between pairs of regions of interest (ROIs), representing the likelihood of a WM path between them, can then be compiled into a connectome. Both histological and DWI-based tract tracing have been qualified as *structural connectivity* (SC; Büchel and Friston, 1997). *Functional connectivity* (FC), on the other hand, refers to the temporal coherence of brain activation, measured by some functional metric such as the blood oxygen-level dependent (BOLD) signal, the basis of functional MRI (fMRI). FC is derived using a purely statistical approach, demonstrating associations between pairs of spatially distinct brain regions, with respect to the signals they generate. It does not typically provide information about causal relationships between regions; i.e., the existence of an axonal projection between them which could transmit action potentials and thus cause the functional association. Connectivity has also been inferred from group-wise correlations of imaging signals and brain morphology. This has been used, for instance, to investigate the correlative structure of grey matter (GM) density (Mechelli et al., 2005), cortical thickness (Lerch et al., 2006), and fluorodeoxyglucose (18F) positron emission tomography (FDG-PET; Friston et al., 1993). Such associations have been posited to arise from mutually trophic influences in the brain (Mechelli et al., 2005) – including but not exclusive to those mediated by direct anatomical connections – as well as genetic factors which govern the coordinated growth of neural tissue in an individual.

The use of any of these modalities to derive or infer connectivity has its own set of limitations. (1.) For CoCoMac, this includes the strong assumption of interspecies homologies, and the large variation in scope and parcellation schemes used in tract tracing studies (Stephan et al., 2000; Kötter and Wanke, 2005). Moreover, as connectivity derived from this database depends exclusively on the available tract tracing literature, its information is incomplete, particularly for contralateral connections. (2.) Probabilistic DWI tractography is biased by the diffusion environment of a particular seed or target point, as well as the distance between them (Dauguet et al., 2007; Mukherjee et al., 2008). As a tract sequentially samples the DWI evidence, it accumulates error, and as a consequence will be biased both by the distance between seed and target points, and the complexity of the WM encountered along the path. For example, the presence of crossing, kissing, or fanning fibres along the streamline (i.e., where the orientation of a single WM fibre has less certainty) will result in a lower probability of that streamline reaching a target. (3.) Resting state (intrinsic) BOLD connectivity (iFC) has been shown to be highly susceptible to both physiological noise and motion artifacts, which create systematic spurious correlations (Power et al., 2012). More conceptually, as functional correlations between brain regions are not sufficient to demonstrate anatomical

connectivity, the interpretation of such associations is ambiguous. (4.) Group-wise correlations in GM morphology provide valuable information about the coordination of growth or atrophy of GM tissue, which may reflect both environmental and genetic processes. However, the degree to which this represents either SC or FC is unclear, and thus would benefit from a comparison to both types of information. Given these limitations, there is a clear need to compare and contrast the connectomes derived from each of these modalities, in a way that might provide insight into these limitations.

The use of alternative terminology for connectivity has led to a confusion over what the term "connectivity" actually refers to, in a given context. While qualifying the term (e.g., as functional, structural, or effective) can reduce this ambiguity, the term is frequently presented in an unqualified way to represent a wide array of meanings. These can range from simple Pearson or partial correlations in functional time series or structural morphology, beta weights in general linear models, group-wise contrasts in DWI-derived measures such as fractional anisotropy or tract-based spatial statistics (TBSS; Smith et al., 2007), streamline counts in DWI-based probabilistic tractography, causal predictions derived from approaches such as structural equation modelling (SEM; McIntosh and Gonzalez-Lima, 1994), Granger causality (Roebroeck et al., 2005), or dynamic causal modelling (DCM; Daunizeau et al., 2011), disconnections caused by lesions or surgical resection, meta-analytic modelling of functional co-activations (MACM; Eickhoff et al., 2010), and actual tract tracing experiments in non-human primates or other animal models. Moreover, connectivity methods typically employ a variety of processing and pre-processing steps which can have a substantial impact on the results and how they are interpreted. For example, in an attempt to remove motion artifacts and physiological noise from the BOLD signal, it is common to regress out the global tissue signal. This processing step transforms a primarily positive distribution of correlation coefficients to a balance of positive and negative ones (Murphy et al., 2009; Schölvinck et al., 2010; Fox et al., 2009). The interpretation of either distribution is clearly different, yet both are referred to as functional connectivity. Thus, while each of these approaches yield unique and scientifically valuable information about how the brain is connected (or disconnected), they each result in quite distinct, and even contradictory, estimates of this phenomenon.

In this article, we start with two main assumptions. The first is that the term "connectivity" refers to a real, physical property of the brain: namely, a summary metric of physical axonal projections originating in a source region A, terminating on synaptic junctions in a (spatially remote) target region B, and mediating the transmission of information (via action potentials) from A to B. To distinguish this property from other commonly employed (and thus meaning-laden) references to connectivity, we will hereafter refer to it as *physical connectivity* (however, we note that the terms "connectivity", "physical connectivity", and "effective connectivity" ought to be considered synonymous). The second is that all of the varied approaches to measuring and modelling connectivity (for instance, those described above) have the ultimate goal of approximating physical connectivity. These assumptions are motivated by the observation that, where connectivity findings are related to biological, behavioural, or pathological phenomena, they are typically interpreted in physical terms (e.g., related to lesion results, known networks in animal models, or putative effective networks in humans). It is our position that such a strict and physically meaningful definition of connectivity is necessary to resolve the ambiguity which frequently accompanies interpretations of connectivity findings, as well as providing a common reference point against which to base comparisons of these findings. On the basis of these two assumptions, we propose that the ultimate goal of any connectomics approach is the ability to model physical connectivity; or, put differently, the ability to make causal inferences about pairs of brain regions, which would be highly informative about the ways in which brain networks process information, and how this breaks down in particular neurological disorders.

While the methods outlined above provide us with a number of distinct windows into physical connectivity, each is limited in the extent to which it can support inferences about it, and each is likely to produce its own unique set of false positives and false negatives. Despite this, it is common – typically due to data availability or a desire for simplicity – to draw conclusions about physical connectivity based solely on observations obtained from a single methodology. It is important to understand, however, how these separate lines of evidence agree (signifying corroboration) and how they disagree (signifying either contradiction or complementarity). For example, evidence of *both* a physical connection and covariance in functional time series provides stronger support for an inference

of physical connectivity between two regions than either finding considered in isolation. In the present study, we therefore characterized the similarities and differences between connectomes derived from four modalities: (1.) the CoCoMac database; (2.) resting state BOLD; (3.) DWI; and (4.) cortical thickness correlations. We compared both weighted and binarized versions of the networks, and used k-means and hierarchical clustering techniques (Bellec et al., 2010; Kelly et al., 2012b) to compare the ways in which these networks form stable clusters, establishing a "consensus" clustering which incorporates evidence from all four modalities. Additionally, we contrasted the results of state-of-the-art neuroimaging data with high spatial, temporal, and angular resolution, with those of an older, lower-resolution dataset, to assess the degree to which intermodal correspondence is affected by these improved techniques. Finally, for the human imaging modalities only, we considered results derived from a functionally-derived, higher granularity parcellation scheme (the Cambridge parcellation; Urchs et al., 2015), in order to assess the degree to which correspondence is dependent on the choice of parcellation.

2. Methods

2.1. Subjects and data

Subjects were obtained from the Enhanced Nathan Klein Institute Rockland Sample (NKI), which is freely available (Nooner et al., 2012). This cohort is the first public release, comprised of 181 participants. After excluding seven subjects on the basis of extreme age (<10 or >80), we analyzed data for 174 subjects aged 10-80 years, and data collection received ethics approval through both the Nathan Klein Institute and Montclair State University. Written informed consent was obtained from all participants, and in the case of minors, also from their legal guardians. All imaging data was acquired from the same scanner (Siemens Magnetom TrioTim, 3.0T). T1-weighted images were obtained using a MPRAGE sequence (TR = 1900 ms; TE = 2.52 ms; voxel size = 1 mm isotropic). Resting state fMRI was performed in three ways: standard (TR = 2500 ms; TE = 30 ms; voxel size = 3 mm isotropic; duration = 5 min); high temporal resolution multiplexed (TR = 645 ms; TE = 30 ms; voxel size = 3 mm isotropic; duration = 10 min); and high spatial resolution multiplexed (TR = 1400 ms; TE = 30 ms; voxel size = 2 mm isotropic; duration = 10 min). DWI was collected with a high spatial and angular resolution (TR = 2400 ms; TE = 85 ms; voxel size = 2 mm isotropic; b = 1500 s/mm^2; 137 gradient directions). Only subjects who passed quality control on all modalities were included in the analyses: specifically, 89 subjects were used for comparison and 92 subjects were rejected for age, data quality, or missing data. Of the subjects used for analysis, age ranged from 11-77 years, with 36 males and 53 females. This same set of subjects was used to perform all subsequent analyses.

To investigate the stability of results across independent cohorts and scanning parameters, we performed an identical analysis with subjects from the International Consortium for Brain Mapping (ICBM) cohort (Mazziotta et al., 2001), comprised of 96 individuals; 47 male, aged 19-85 years. The protocol was approved by the Research Ethics Committee of the Montreal Neurological Institute and Hospital, and each subject provided informed consent. No subjects had a history of neurological or psychiatric disorders. For this cohort, T1-weighted images were obtained using a Siemens Sonata 1.5T MR scanner, with a 3D gradient echo sequence (TR=22 ms; TE=9.2 ms; voxel size = 1 mm isotropic). DW images were obtained from the same Siemens scanner as for the T1 images, using a single-shot echo planar sequence, with sensitivity encoding and a parallel imaging factor of 2.0. 60 axial slices were acquired with a thickness of 2.5 mm and no interslice gap (TR = 8000 ms; TE = 94 ms; b=1000 s/mm^2; 30 gradient directions). Three repetitions of the sequence were collected in order to achieve a sufficient signal to noise ratio. 138 volumes of resting-state BOLD signals were recorded on the same scanner using a 2D echoplanar BOLD MOSAIC sequence (TR=2000 ms; TE=50 ms; voxel size = 4 mm isotropic; duration = 4.6 min).

CoCoMac data was obtained from $\sim$ 300 tract tracing studies, and $\sim$ 150 articles describing brain atlases (Bakker et al., 2012). Data was entered manually, according to rules designed to facilitate their standardization for use as metadata (Stephan et al., 2001; Kötter, 2004).

2.2. Cortical parcellations

Cortical ROIs were based on the Regional Map (RM) defined by Kötter and Wanke (2005), which were selected using structural, functional, and topographic criteria with the express purpose of maximizing homology across primate species, including humans. The RM parcellation consists of 80 cortical regions (see Table 1 and Figure 1) which were hand-drawn on the macaque F99 template surface (Van Essen, 2002, 2004) under supervision of its creator, Rolf Kötter (Bezgin et al., 2012, 2008). Using Caret software, the parcellation was then transformed onto the PALS human template (Van Essen, 2005) by means of a landmark-based deformation defined by a set of major sulci and gyri, along with functional activation patterns, considered homologous between the two species (Van Essen and Dierker, 2007). The latter was transformed to the ICBM population template surface (Lyttelton et al., 2007), for the cortical thickness analysis, and projected into volumetric space along the surface normals (Bojak et al., 2011), using ModelGUI software (`http://www.modelgui.org`), for the fMRI analysis (σ_N=3 mm, σ_T=1.67 mm); and in-house scripts, for the DWI analysis.

In addition to the RM parcellation, we used a second, functionally-derived parcellation scheme as a means of assessing the effect of granularity and parcel derivation on the cross-modal correspondence. This so-called "Cambridge parcellation" was derived as follows (Urchs et al., 2015). Resting-state fMRI data were obtained from the Cambridge sample of approximately 200 young healthy adults, as part of the 1000 Functional Connectomes Project (Liu et al., 2009). Parcellations were derived using a method called bootstrap analysis of stable clusters (BASC; Bellec et al., 2010), across multiple scales (7, 12, 20, 36, 64, 122, 197, 325, and 444), determined using a data-driven method called MSTEPS (Bellec, 2013). For the current study, we utilized the 122-region (per hemisphere) parcellation. An important difference between the original Cambridge and RM parcellations was that functional parcels were composed of voxels which were not necessarily spatially contiguous, often including homotopic regions in both brain hemispheres. To facilitate the comparison between the two parcellation approaches, the functional parcels were further subdivided into spatially contiguous subparcels that spanned only one hemisphere. This volumetric parcellation was then transferred to the ICBM-152 asymmetric template surface using ModelGUI software; specifically, the mode of all non-zero voxel values within a 10 mm radius of a given vertex was assigned to that vertex. Since only regions which overlapped with the cortical surface are comparable to cortical thickness estimates, all subcortical regions were excluded in this process. The final number of regions was thus 168 (84 per hemisphere). The final surface-based Cambridge parcellation (right hemisphere) is shown in Figure 1.

2.3. Terminology

In this study, the graph theoretical term “edges” is used to refer to inferred connectivity and “vertices” is used to refer to individual ROIs. An edge is associated with a “weight”, which refers to the strength of the relationship between the ROIs it connects; the assignment of edge weights is described in the proceeding sections.

2.4. CoCoMac network

Projections between RM regions were compiled from all CoCoMac literature entries using the Objective Relational Transformation method (ORT; Stephan et al., 2000), which integrates multiple source brain maps – of different scales and covering different parts of the brain – using relations based upon textual statements in the literature (“containing”, “contained by”, “identical to” “disjoint from”, etc.). Weights for these projections were obtained using the relative density values, as reported in the literature, which represent the density of tracer substance found in a labelled site. Where experimental evidence existed, density values range from 0 (no label) to 3 (dense label); where label was present but no density information was provided in the article, we have here assigned a value of 2, presuming this to be the “expected value” for a label of unknown density. To make the CoCoMac database comparable with the imaging-based networks, which are undirected, we made it undirected by assigning:

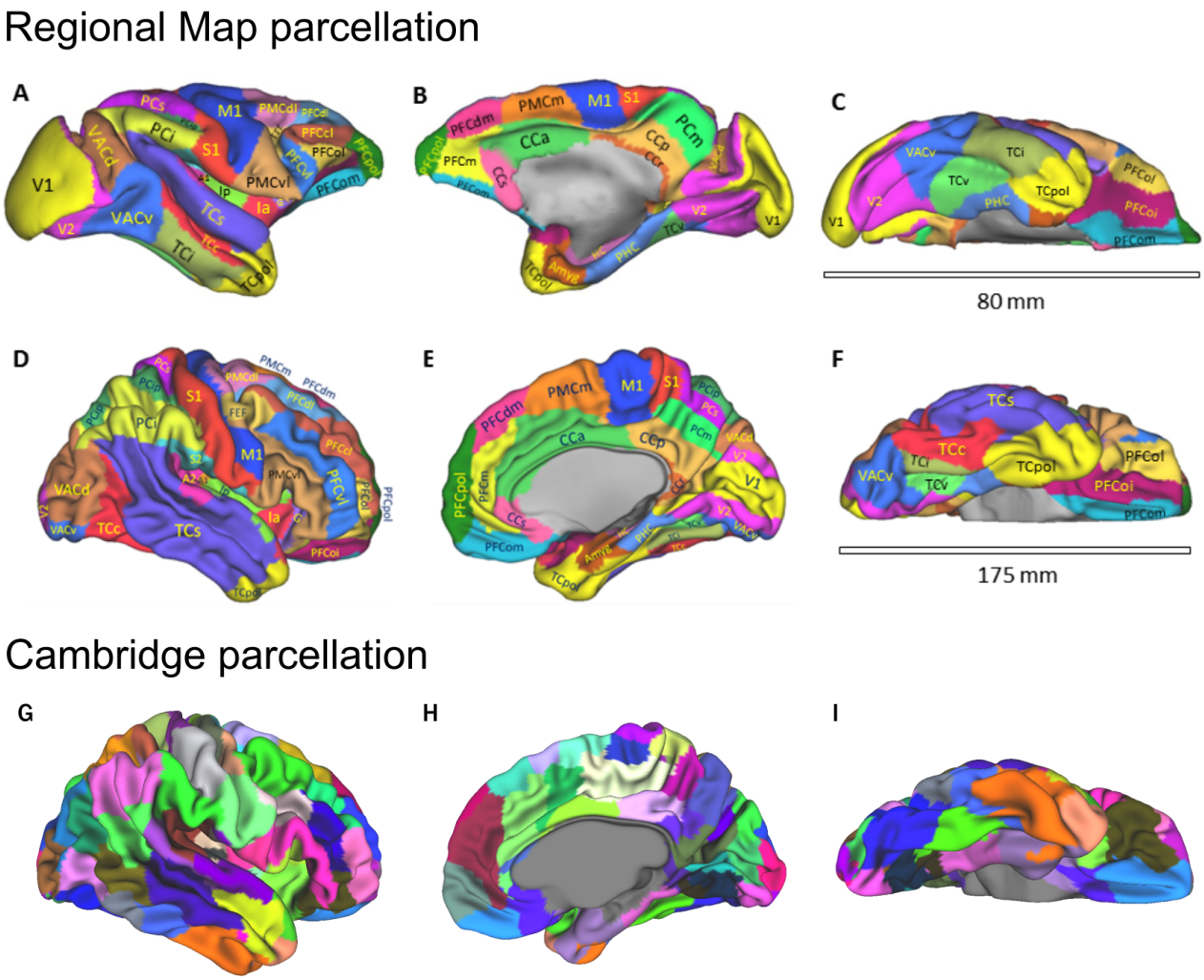


Figure 1: **Cortical parcellations.** At top, the 80-region Regional Map (RM) parcellation (right hemisphere only), shown on the macaque F99 template surface (A,B, and C), and on the human ICBM-152 template surface (D, E, and F). A list of abbreviations is provided in Table 1. Adapted with permission from Bezgin et al. (2012). The bottom row shows the 168-region functionally-derived Cambridge parcellation (G,H, and I), used to perform complementary human-only comparisons, on the same surface.

(1.) the sum of density values over both directions, if evidence existed for both; or (2.) the density value for the known direction, if there was experimental evidence for one direction but not the other. This approach was taken under the assumption that a bidirectional projection is likely to result in stronger associational coupling than a unidirectional one. This resulted in an 80x80 matrix R_{cocomac}.

2.5. Cortical thickness network

Cortical thickness was computed using the CIVET pipeline (Zijdenbos et al., 2002). Briefly, this involves an initial N3 correction for field non-uniformities (Sled et al., 1998), linear and nonlinear registration to the MNI152

Table 1: Regional Map Parcellation Names

Abbreviation	Lobe	Full name
A1	Temporal	Primary auditory cortex
A2	Temporal	Secondary auditory cortex
Ia	Temporal	Anterior insula
Ip	Temporal	Posterior insula
Amyg	Limbic	Amygdala
CCa	Limbic	Anterior cingulate cortex
CCp	Limbic	Posterior cingulate cortex
CCr	Limbic	Retrosplenial cingulate cortex
CCs	Limbic	Subgenual cingulate cortex
FEF	Frontal	Frontal eye field
HC	Limbic	Hippocampus
M1	Frontal	Primary motor cortex
PFCcl	Frontal	Centrolateral prefrontal cortex
PFCdl	Frontal	Dorsolateral prefrontal cortex
PFCdm	Frontal	Dorsomedial prefrontal cortex
PFCm	Frontal	Medial prefrontal cortex
PFCoi	Frontal	Intermediate orbital prefrontal cortex
PFCol	Frontal	Orbitolateral prefrontal cortex
PFCom	Frontal	Orbitomedial prefrontal cortex
PFCvl	Frontal	Ventrolateral prefrontal cortex
PFCpol	Frontal	Polar prefrontal cortex
PHC	Limbic	Parahippocampal cortex
PMCdl	Frontal	Dorsolateral premotor cortex
PMCm	Frontal	Medial (supplementary) premotor cortex
PMCvl	Frontal	Ventrolateral premotor cortex
S1	Parietal	Primary somatosensory cortex
S2	Parietal	Secondary somatosensory cortex
PCi	Parietal	Inferior parietal cortex
PCip	Parietal	Cortex of the intraparietal sulcus
PCm	Parietal	Medial parietal cortex
PCs	Parietal	Superior parietal cortex
TCc	Occipital	Central temporal cortex
TCi	Occipital	Inferior temporal cortex
TCs	Temporal	Superior temporal cortex
TCpol	Temporal	Polar temporal cortex
TCv	Temporal	Ventral temporal cortex
V1	Occipital	Primary visual cortex
V2	Occipital	Secondary visual cortex
VACd	Occipital	Anterior visual cortex, dorsal part
VACv	Occipital	Anterior visual cortex, ventral part

population template (Mazziotta et al., 2001), tissue classification with partial-volume estimation (Tohka et al., 2004), and the approximation of GM/WM and GM/CSF boundaries using the Constrained Laplacian Anatomic Segmentation using Proximity (CLASP) algorithm (MacDonald et al., 2000; Kim et al., 2005). This results in surface meshes with 40,962 vertices per hemisphere, which are subsequently coregistered to a population average surface template, such that each vertex is in correspondence across all subjects (Lyttelton et al., 2007). Quality control was performed by evaluating each subject visually, using ModelGUI software (`http://www.modelgui.org`);

45 subjects were rejected on the basis of having substantial issues, due variously to: (1) motion artifacts, (2) masking errors, (3) misclassification of white matter leading to modelling failure, or (4) failure of the normalization step. The severity of these issues was ranked on a scale of 0 (no obvious issues), 1 (minor issues), 2 (at least one major issue), or 3 (severe issues). All subjects with scores of 2 or 3 were excluded from the analysis. Thickness values were smoothed with a 5 mm Gaussian kernel, corrected for the effects of age and sex, and then averaged for each ROI in the RM parcellation. Correlations were computed across all individuals for each pair of ROIs, resulting in an 80x80 correlation matrix $R_{\text{thickness}}$, which was then thresholded for family-wise error (FWE) using a false discovery rate (FDR; Benjamini and Hochberg, 1995) of $q < 0.05$.

2.6. DWI network

The diffusion-weighted images (DWI) were converted to 4D volumes, and cleaned of motion and other artifacts using DTIPrep (Oguz et al., 2014). DTIPrep corrects artifacts where possible, and excludes directions from the data when correction is not possible. Where artifact rejection resulted in fewer than 110 retained directions, the dataset was excluded from the analysis. Of the subjects which passed quality control for other modalities, two additional subjects were excluded on this basis (see Figure S7). The cleaned 4D diffusion volumes were then structurally unwarped: the T1 volume was linearly registered to the average B_0 volume; the average B_0 volume was non-linearly registered to the overlayed T1 volume; and the non-linear transform was applied to each gradient volume. The average B_0 for the cleaned unwarped diffusion data was then linearly registered to the T1 volume in stereotaxic space, and the resultant transform applied to the gradient volumes, and the rotational component to the directional vectors. This yielded a cleaned unwarped 4D diffusion volume in stereotaxic space. The scaling component was then removed from the transform, inverted, and applied to the diffusion gradients to produce a cleaned unwarped 4D diffusion volume at native scale. Both 4D diffusion volumes were then processed with FSL's *bedpostx* to determine the distribution of diffusion directions at each voxel. Probabilistic tractography was then performed using FSL's *probtrackx* package (Behrens et al., 2007), with WM as a seed mask, surrounding voxels (including subcortical GM) as a stop mask, and cortical and subcortical GM as a target mask. Each WM voxel was seeded from 10,000 random positions. Tractography was performed using the *bedpostx* results for the stereotaxic diffusion volume with options to produce a distance-bias corrected estimate of the number of connections between any two voxels in the target mask; tractography was performed using the *bedpostx* results for the native-scale diffusion volume with options to produce the average length of the connections between any two voxels in the target mask. These results were then compiled for both atlases (RM and Cambridge) to produce the total number of connections between pairs of ROIs, and the average length of those connections. An additional correction for ROI size was performed by dividing these distance-corrected values by the product of the surface areas of the two ROIs:

$$T_{A,B}^{\text{adj2}} = \left(\frac{T_{A,B}^{\text{adj1}}}{SA_A \cdot SA_B} \right)^{0.1} \tag{1}$$

where $T_{A,B}^{\text{adj1}}$ denotes the distance-corrected tract count, $T_{A,B}^{\text{adj2}}$ the distance- and size-corrected tract count, and SA_A the surface area of ROI A. As the resulting distribution remained highly skewed, an exponential transformation of 0.1 was applied to produce a more normal distribution. Finally, adjusted matrices for each subject were averaged across the group to obtain a single 80x80 matrix R_{dwi}.

2.7. fMRI network

For the present analyses, we used the high temporal resolution (TR=645ms) multiband fMRI data of the NKI sample. The datasets were preprocessed using the Neuroimaging Analysis Kit, version "ammo" 0.7.1 (NIAK; Bellec

et al., 2011), under CentOS release 5.5 with Octave version 3.6.2 and the Minc toolkit version 0.3.18. All analyses were executed using the Pipeline System for Octave and Matlab version 1.0.2 (PSOM; Bellec et al., 2012). The parameters of a rigid body motion were first estimated at each time frame of the fMRI dataset, using the median of the run as a target. This target fMRI volume was coregistered with a T1 scan of the same individual. The rigid-body transform, fMRI-to-T1 transform and T1-to-stereotaxic transform were all combined, and the functional volumes were resampled in the MNI space at a 3 mm isotropic resolution. The "scrubbing" method of Power et al. (2012) was used to remove the volumes with excessive motion (frame displacement greater than 0.5). The following nuisance parameters were regressed out from the time series at each voxel: slow time drifts (basis of discrete cosines with a 0.01 Hz high-pass cut-off), average signals in conservative masks of the white matter and the lateral ventricles as well as the first principal components (95% energy) of the six rigid-body motion parameters and their squares (Giove et al., 2009). Manual individual review of the raw fMRI volume, fMRI-T1 and T1-stereotaxic coregistration lead to the identification of a major issue (i.e. signal loss in large portions of the fMRI volumes) in one subject, and minor issues in another set of 44 subjects (i.e., one or multiple issues amongst bad coregistration of the central sulcus, cingulate gyrus, lateral ventricles, temporal lobes, hippocampus, parieto-occipital sulcus, or bad brain masking). This preprocessing quality control process is documented at `https://github.com/SIMEXP/niak_manual/raw/master/qc_manual_v1.0/qc_manual_niak.pdf`. For this first-pass quality control, since we were aiming for the best possible quality data, we rejected all subjects with either major or minor issues (45 total). Of the remaining 126 subjects, an average of 65.3 of 900 frames per subject, or 7%, were scrubbed due to the presence of motion artifacts. As a second-pass quality control, we rejected subjects with less than 450 frames retained (see Figure S7). This threshold corresponds to $\sim$ 5 min of signal, which is sufficient to produce stable correlations (Van Dijk et al., 2010). This resulted in two further subjects being rejected from subsequent analysis. Finally, as an alternative preprocessing step, a correction for global signal fluctuations was implemented by regressing out the first principal component of the time series of all voxels falling into a brain mask segmented from functional images, as described in Carbonell et al. (2011). For each subject, correlations were computed between the average time series for each pair of RM regions. The final 80x80 matrices, R_{fmri} and $R_{\text{fmri}-\text{gsr}}$, were obtained by computing the average correlation coefficients across subjects.

2.8. Default mode network

We defined the default mode network (DMN) to roughly match those regions selected by Greicius et al. (2009). The ROIs included in this network were: medial parietal cortex (PCm), posterior cingulate cortex (CCp), retrospenial cingulate cortex (CCr), ventral temporal cortex (TCv), parahippocampal gyrus (PHC), hippocampus (HC), medial prefrontal cortex (PFCm). Notably, this network includes only medial regions, and thus excludes the temporoparietal junction, although this region is also commonly considered a component of the DMN.

2.9. Resampling

Resampling of the neuroimaging data was performed in order to cross-validate all findings from the analyses described below. This allows us to assess the degree of variance in connectivity scores obtained from different subsamples of the cohort and, conversely, the relative stability of individual modalities and the correspondence between them. After quality control of all three modalities, 89 subjects were used to generate samples. This consisted of generating 500 random subsamples, without replacement, of size $B = \frac{N}{2}$ (where N is the number of subjects). This was done for each neuroimaging modality and connectivity matrices were computed for each subsample, as described in the preceding sections. For the thickness network, we generated group-wise correlations for each subsample. For the fMRI and DWI networks, we obtained the average of subject-wise correlations and tract counts, respectively, across the subsample. This approach was applied to all comparisons described below. R_{cocomac} was not resampled, as it is based upon a single connectivity matrix.

2.10. Weighted comparison

To compare connectivity matrices for each modality we first correlated each with one another, to obtain six Spearman rank coefficients representing each pairwise correlation. The Spearman rank coefficient was used since, for the CoCoMac database, the distribution was too skewed to justify a Pearson approach. In the case of R_{cocomac}, only edges for which experimental evidence existed were used to compute this coefficient. We performed the same analysis on: (1.) intra- and interhemispheric connections; (2.) ipsilateral connections within each lobe only; and (3.) connections with low (14-110 mm), medium (110-160 mm), and high (160-270mm) WM projection distances (these ranges were chosen to partition the edges into three groups of equal size). Projection distance was computed from the DWI tracts, using the *probtrackx* tool, as described above (Behrens et al., 2007). Finally, in order to get a measure of the within-modality reliability for each of the human imaging modalities, we performed a split-half analysis, wherein one half of the sample was compared to the other, using this same weighted comparison. This analysis was performed for 250 random split-half assignments.

2.11. Weight-thresholded binarized comparison

Since the assignment of edge weight in each modality may not have a good correspondence, we next compared binarized versions of these matrices, assigning either one or zero to each element, and thresholding at a density of 58%, corresponding to the maximal density of R_{cocomac} (i.e., considering only non-zero edges). In some cases (for example, the human-only comparison illustrated in Figure 6), analyses were also performed at all other of the six possible densities of R_{cocomac}. Comparisons were performed by obtaining an accuracy score for each pair of modalities $[f, g]$, where accuracy is equal to the ratio of all matches to all possible ROI pairs $[i, j]$, N_{pairs}:

$$\Phi_h = \frac{\sum_{i=1}^{N_{pairs}} \sum_{j=i+1}^{N_{pairs}} \left(TP_h^{(ij)} + TN_h^{(ij)} \right)}{N_{\text{pairs}}} \tag{2}$$

where h denotes a comparison between two matrices, $TP^{(ij)}$ denotes a true positive ($R_f^{(ij)} = R_g^{(ij)} = 1$), and $TN^{(ij)}$ denotes a true negative ($R_f^{(ij)} = R_g^{(ij)} = 0$). Notably, "truth" in this context refers to a case where two matrices agree. However, any two random matrices will agree by chance, at a rate determined by their densities. To account for this random agreement, we further define an adjusted accuracy score ACC, by normalizing between the value expected by random chance, Φ_{rand}, and 1:

$$ACC = \frac{\Phi - \Phi_{\text{rand}}}{1 - \Phi_{\text{rand}}} \tag{3}$$

The derivation of Φ_{rand} is given in the Supplemental Material. The resulting value varies from 0 to 1, where 1 indicates perfect agreement, and 0 indicates agreement no better than that expected by chance. For comparisons involving R_{cocomac}, only edges for which experimental evidence existed were used to compute ACC.

To obtain a score integrating information across all modalities, we derived the "correspondence index", CI, which is similarly the ratio of matches for each ROI pair $[i, j]$ across all M modalities, to the total number of comparisons, $H = M(M-1)/2 = 6$:

$$CI^{(ij)} = \frac{1}{H} \sum_{h=1}^{H} (TP_h^{(ij)} - TN_h^{(ij)}) \tag{4}$$

Where edges in R_{cocomac} had no experimental evidence, only the other three modalities were used to compute CI. Note that the matrix obtained from the difference $TP_h^{(ij)} - TN_h^{(ij)}$ will have a value of +1 for a true positive, -1 for a true negative, and 0 for a false positive or false negative; thus the sign of CI indicates the type of agreement and the magnitude indicates the degree of intermodal correspondence. The inverse of the absolute value of this score $(1 - |CI|)$ represents the degree of intermodal disagreement.

As for the weighted comparison, the same 250 split-half assignments were used to test within-modality reliability for each of the human imaging modalities, using these weight-threshold binary comparisons.

2.12. K-means and hierarchical clustering

Clustering approaches can be applied to network graphs in order to obtain groups of vertices which are maximally connected. Connectivity matrices for each modality were considered "similarity" matrices for the purpose of hierarchical and k-means clustering. Following Bellec et al. (2010) and Kelly et al. (2012b), each matrix was clustered into k = 2, 4, 6, 8, 10, 12, 14, 16, 18, and 20 partitions, across all subsamples. For hierarchical clustering, we used the NIAK library (Bellec et al., 2011, 2012) and Ward's criterion (Ward, 1963). For k-means clustering, we used the *kmeans* function from MATLAB R2014a (©The MathWorks, Inc., 2014). For each subsample b, each modality m, and each partition size k, we computed an adjacency matrix $A^{(m,b,k)}$, such that each element $a_{ij}^{(m,b,k)} = 1$ if ROIs i and j were in the same cluster, and 0 otherwise. A within-modality *stability matrix* $S^{(m)}$, describing the proportion of subsamples in which a pair of ROIs was in the same cluster *within a modality*, was then constructed for each as:

$$S^{(m,k)} = \frac{1}{B} \sum_{b=1}^{B} A^{(m,b,k)} \tag{5}$$

where B is the total number of subsamples.

Similarly, a cross-modality *consensus matrix* C, describing the proportion of subsamples in which ROIs i and j were in the same cluster *across modalities*, was constructed as the average agreement between each pair of modalities:

$$C^{(k)} = \frac{1}{H} \frac{1}{B} \sum_{b=1}^{B} \sum_{h=1}^{H} A^{(h,b,k)} \tag{6}$$

where h represents one of six pairwise comparisons between modalities, and $A^{(h,b,k)}$ is the adjacency matrix such that each element $a_{ij}^{(h,b,k)} = 1$ if ROI i from one modality is in the same cluster as ROI j from the other, and 0 otherwise. This formulation allows us to quantify the degree to which similarity information from all four modalities converge on a specific clustering solution (see Bellec et al., 2010).

Since it was not subsampled, the same instance of R_{cocomac} was clustered 500 times. For hierarchical clustering, because this matrix contained only six possible integer values, and was highly skewed towards higher values, the clustering routine encountered a high number of ambiguous choices; i.e., cases where the similarity of two possible clusters was identical. In these cases, the choice was made pseudorandomly; thus, the hierarchical stability matrix of R_{cocomac} largely captured algorithmic (rather than population) stability (see Discussion).

Finally, to determine the efficacy of a given clustering solution, we defined a *stability contrast* σ (Bellec et al., 2010), which quantifies the within-cluster stability of an ROI i relative to its maximal stability within the other clusters (between-cluster stability). Values of σ can range from -1 (greater between- than within-cluster stability) to +1 (greater within- than between-cluster stability).

2.13. Signal-to-noise

The signal-to-noise ratio (SNR) provides a measure of how well a target signal (e.g., a connectivity score) can be distinguished from its noise component. When comparing connectivity inferred from multiple modalities, it is informative to assess the degree to which intermodal correspondence is associated with SNR; in other words, how does the variability of a connectivity score relate to the degree to which it corresponds across modalities? For each human-derived modality m and ROI i, we assessed SNR with respect to the variance in connectivity scores across subsamples, as the ratio of the mean score μ_m to its standard deviation σ_m:

$$SNR_m^{(i)} = \frac{\mu_m^{(i)}}{\sigma_m^{(i)}}$$

SNR for a given ROI pair $[i, j]$ was then computed as the average of SNR_m^i and SNR_m^j.

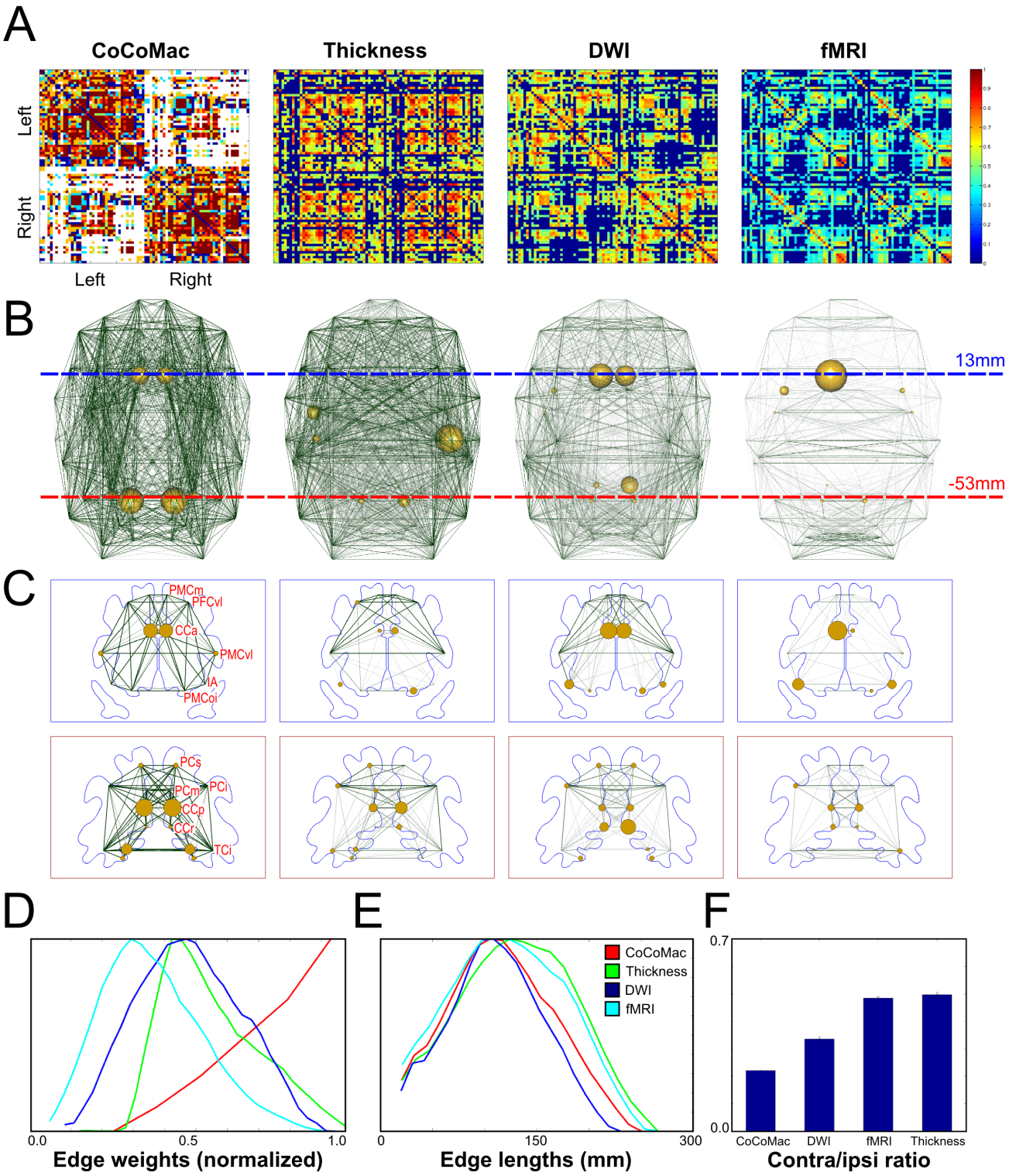
For the CoCoMac database, there is no straightforward means of determining a SNR. Instead, a measure of data completeness, DC was determined by enumerating the experimental injection or labelled sites from which values were determined for each ROI. These values indicate the amount of evidence supporting an inference of connectivity for each RM region (for a detailed description, see Bezgin et al., 2012). Similar to the SNR for human modalities, the average of DC_m^i and DC_m^j was used to represent the SNR for each ROI pair.

SNR scores for each modality, and the mean across modalities, were compared to the CI score, using a Spearman rank correlation. This allowed us to assess the degree to which correspondence could be explained by the amount of noise in the connectivity measures used to compute it.

3. Results

3.1. The networks

Of all possible ROI pairs in R_{cocomac}, 32.7% lack experimental data to confirm either their existence or lack thereof. The majority of these unknown connections are contralateral (58.7% versus 6.9% for ipsilateral connections). 57.6% of possible connections are experimentally confirmed to exist, while only 9.7% are confirmed not to exist. Ipsilaterally, CoCoMac is 80.8% connected, while contralaterally it is only 33.8% connected, excepting unknown connections. Matrices for all networks are shown in Figure 2A, where the unknown CoCoMac edges are shown in white. Qualitatively, these matrices have a number of similarities – particularly in the prominence of homotopic connections (the off-diagonal in the upper-right quadrant), which is strongest for $R_{\text{thickness}}$ and R_{fmri}; as well as the similar contralateral structure of R_{cocomac} and R_{dwi} (if we count the unknowns as nonexistent). This similarity in ipsilateral and contralateral structure is also evident from the superior aspect 3D renderings shown in Figure 2B, as well as the coronal sectional renderings shown in Figure 2C. In this illustration, vertices have been scaled by their distance-weighted betweenness centrality. Notably, all such "hub" vertices are distributed more symmetrically for R_{cocomac} and R_{dwi}, with anterior and posterior cingulate cortex being prominent hubs. Figure 2D illustrates



the edge weight distribution of each network. While the three imaging modalities produce networks with normally distributed values, R_{cocomac} is highly right-skewed, due to its predominance of high density connection weights and

Figure 2: **Cortical networks for each modality**, derived from the NKI dataset, thresholded at the maximal CoCoMac density of 58%. **A.** Normalized connectivity matrices, coloured by edge weight. Note that the colour scale is from 0 to 1, and that blue here indicates low or zero connectivity, as opposed to the fMRI convention of blue representing negative covariance. For R_{cocomac}, white signifies that no tract tracing evidence exists in the CoCoMac database for that ROI pair. **B.** The networks rendered as 3D graphs, from the superior aspect. The darkness and opacity of an edge signifies its weight, while the size of a vertex signifies its distance-weighted betweenness centrality (a measure of hubbiness) in the network. **C.** The same networks rendered on 2D coronal sections, which are shown in the 3D renderings. Sections are located at +13 mm (blue) and -53 mm (red) in MNI stereotaxic space. As for B, edge weight is signified by the darkness, opacity, and size of the line, while distance-weighted betweenness centrality is signified by the vertex size. ROI labels are provided in the leftmost section (see Table 1 for a complete list). **D.** Histogram of edge weights for each modality; note the highly skewed distribution of R_{cocomac}. **E.** Histogram of edge lengths for each modality, using the centroids of RM regions in (human) MNI coordinate space. **F.** The ratio of contralateral to ipsilateral connections for each modality.

scarcity of low-density weights. Figure 2E shows the distribution of edge lengths for each network; $R_{\text{thickness}}$ and R_{fmri} have longer edges on average, which is likely due to their having a greater number of contralateral connections. Accordingly, the ratio of contralateral to ipsilateral connections is low for R_{cocomac} (~ 0.2), intermediate for R_{dwi} (~ 0.3), and high for $R_{\text{thickness}}$ and R_{fmri} (~ 0.5; see Figure 2F). These patterns were virtually unchanged when computed for a reduced subset of the NKI dataset, which included only subjects in the age range 18-40 (77 total), indicating that the larger age range did not have a substantial effect on the connectivity results.

3.2. Weighted comparison

Directly correlating network edge weights between modalities provides a set of simple pairwise comparisons. Correlations between the four weighted networks are shown in Figure 3A. For the NKI dataset, correlation was highest between R_{dwi} and R_{fmri} ($r = 0.64$, $r^2 = 0.41$), and lowest between R_{cocomac} and both $R_{\text{thickness}}$ ($r = 0.23$, $r^2 = 0.05$) and R_{fmri} ($r = 0.27$, $r^2 = 0.07$). By comparison, the ICBM dataset had a very similar pattern, but generally lower correlations - this difference was most prominent between the imaging modalities, and was actually reversed for $R_{cocomac}/R_{dwi}$. When grouping NKI correlations by edge distance (Figure 3B), correlations between the three imaging modalities was highest for short (<110 mm) and medium-length (110-160 mm) connections, and decreased for long (>160 mm) connections. This pattern was similar for ICBM (Figure S1). Figure 3C shows NKI correlations obtained within lobes, ipsilaterally. This results in a general substantial increase in most correlation values, including comparisons for R_{cocomac}. The correlations were heterogeneous, with higher association between the imaging modalities in frontal, parietal, and occipital lobes, and for comparisons involving R_{cocomac} most markedly in the parietal lobe. Correlations were generally comparable between ipsilateral and contralateral connections, with the exception of the $R_{\text{thickness}}$ and R_{fmri} comparison (Figure 3D); this was again similar for the ICBM dataset (Figure S1).

Figure 6A shows intermodal correlations for human-only modalities, obtained from both parcellation schemes, as well as a comparison between fMRI with and without global signal correction (GSR). For all comparisons, connectivity estimates for RM had a higher agreement than for the functionally-derived Cambridge parcellation. Additionally, for both parcellations, fMRI estimates produced with GSR had worse correspondence than without. fMRI estimates with and without GSR themselves had an agreement (r^2) of 0.65 to 0.70, indicating unshared variance of at least 30%. It is also notable that $R_{\text{thickness}}$ had a particularly poor within-modality reliability for weighted comparisons, having an r^2 of less that 0.6, when evaluated using split-half comparisons (Figure S8).

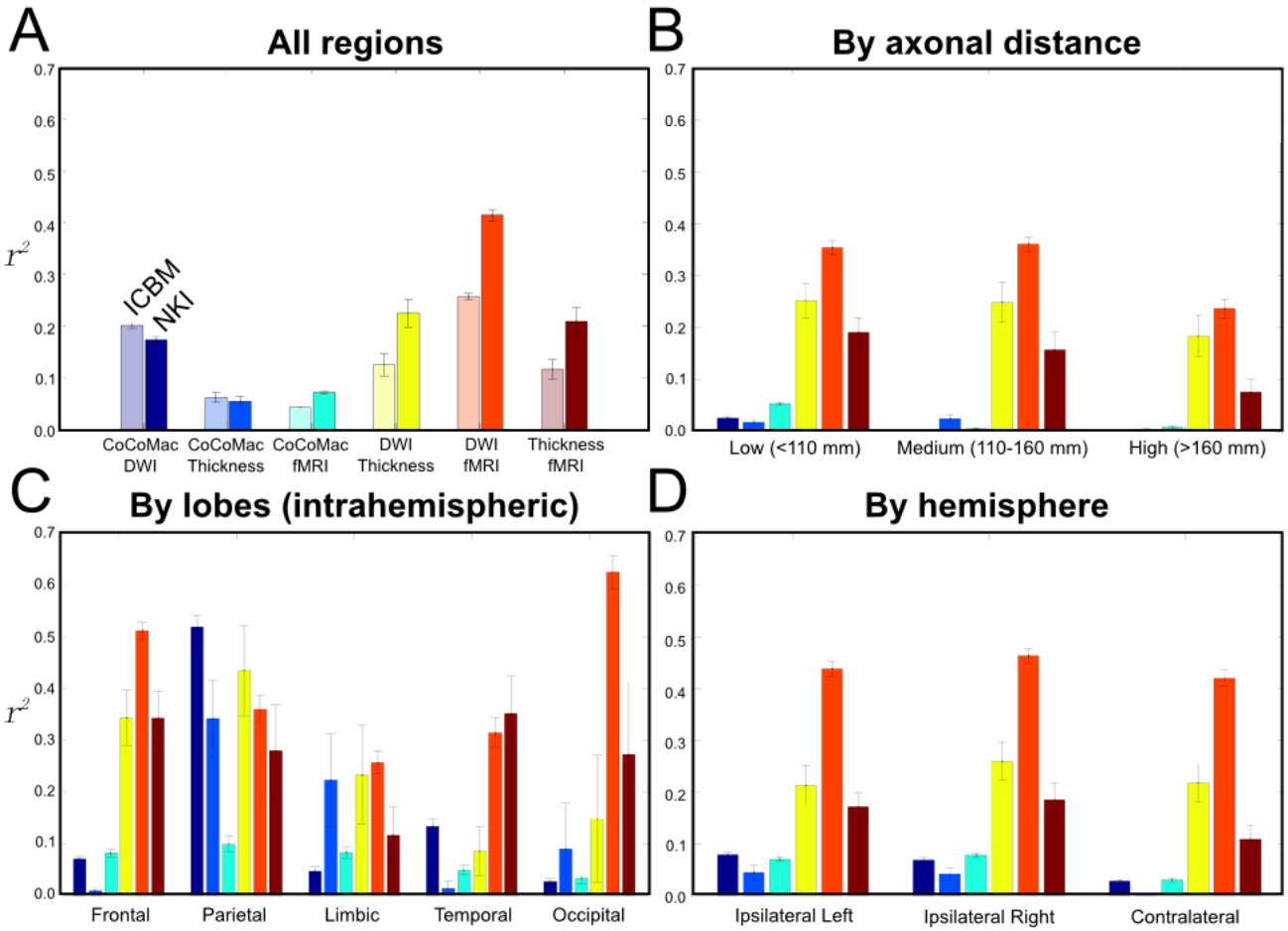


Figure 3: **Weighted comparisons. A.** R^2 values for correlations between each pair of modalities. Results for the ICBM dataset are shown in lighter colours and for the NKI dataset in darker colours (as labelled). Error bars show standard deviation across subsamples. All other panels show the NKI results only, with the same colour coding: **B.** R^2 values for correlations on edges having low, medium, and high axonal distance, measured by DWI tractography. **C.** R^2 values for correlations of ipsilateral edges within lobes. **D.** R^2 values for correlations for right and left ipsilateral connections and contralateral connections.

3.3. Binarized comparison

A major limitation of correlative comparisons is the different ways in which edge weights are derived from each modality. Even given the (strong) assumption that these relationships are monotonic (i.e., they are ordered the same), it is still possible that they are nonlinear, and thus poorly captured through a correlation. Of particular concern is the non-Gaussian, highly skewed distribution of R_{cocomac} (Figure 2D). To address this, we performed a second comparison by binarizing each network to match the maximal density of R_{cocomac} (58%). Figure 4A shows the accuracy (ACC, i.e., the degree to which one network matches the other) for each cross-modal comparison. In general, ACC was fairly poor for both NKI and ICBM datasets (note that the value of ACC varies on the range [0,1], signifying a continuum from random chance to perfect agreement). ACC was higher for NKI, for comparisons between the human imaging modalities, and for ICBM, for comparisons involving $R_{cocomac}$. Notably, relative ACC for comparisons involving R_{cocomac} were improved from the correlative approach, for both datasets – ACC between R_{cocomac} and R_{dwi} was particularly improved. For the imaging comparisons, the patten of ACC was comparable

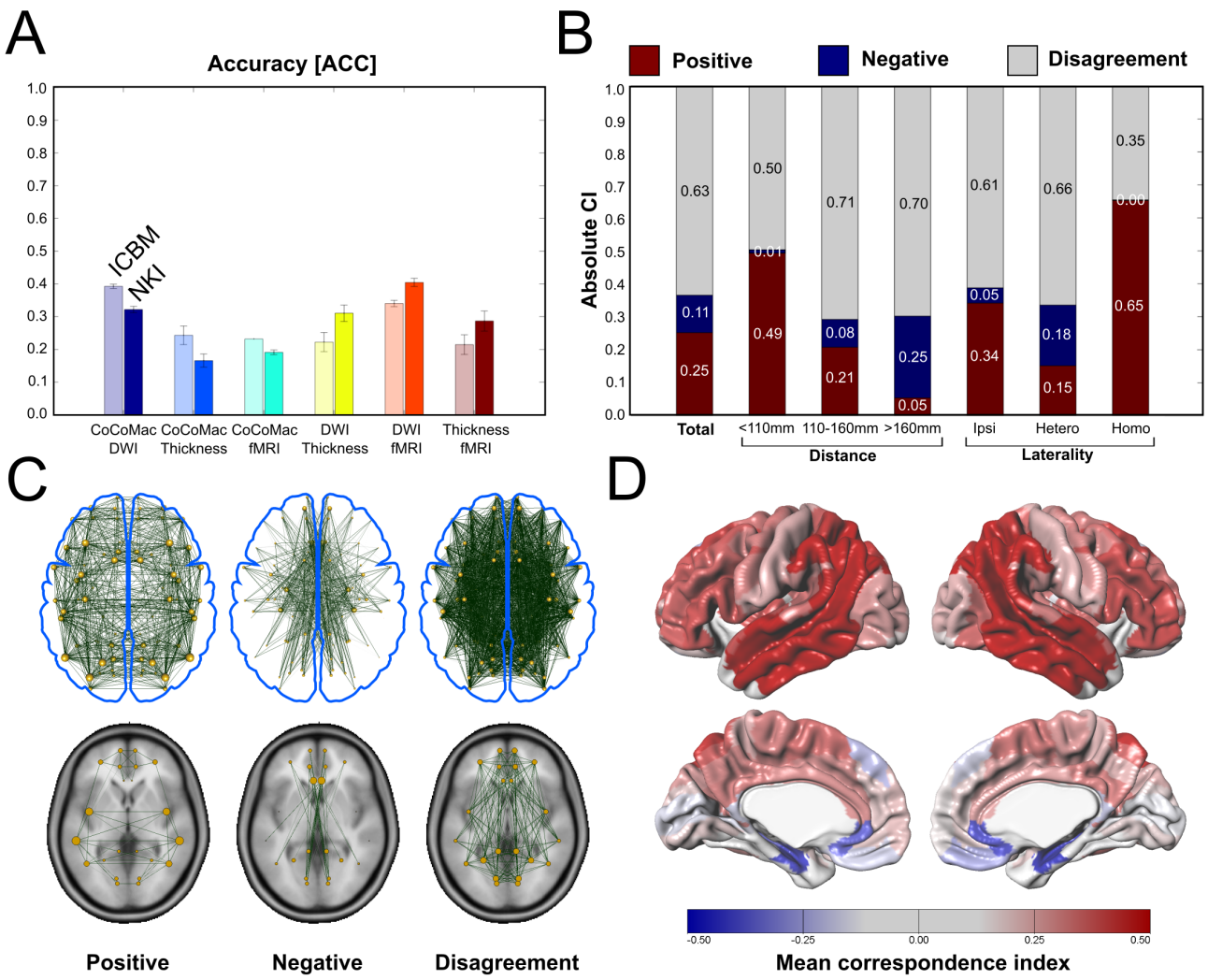


Figure 4: **Binarized comparisons. A.** Accuracy (ACC) for comparisons between each pair of modalities. Note that ACC varies from 0 to 1, such that $ACC = 0$ represents the level of agreement expected by random chance, and $ACC = 1$ represents perfect agreement. Results for the ICBM dataset are shown in lighter colours and for the NKI dataset in darker colours (as labelled). Error bars show the standard deviation across the subsample. Because ACC is normalized between that expected by random chance (0) and perfect agreement (1), its value indicates how much better the agreement between each pair of modalities is from chance. All other panels show the NKI results only: **B.** The correspondence index (CI), which measures cross-modal agreement. Positive CI (red) measures the degree to which all modalities agree that an edge exists, and negative CI (blue) that it does not exist. Disagreement (grey) is the absolute inverse of this value. CI is also shown for low, medium, and high edge distance, for ipsilateral connections, and for heterotopic and homotopic contralateral connections. **C.** Positive and negative CI and disagreement, shown from the superior aspect as a 3D graph (top) and as horizontal 2D sections (bottom; 0 mm in MNI stereotaxic coordinates, with ±10 mm clipping bounds; superimposed on the ICBM-152 linear template image). Edge darkness and opacity signify the strength of each index, and the size of the vertex signifies the sum for that ROI. **D.** Mean CI painted on the ICBM template surface. Lateral, superior, and cingulate ROIs show a higher mean positive CI (red), while medial frontal and temporal gyri show a higher mean negative CI.

to that of the correlative results, with R_{dwi} and R_{fmri} having the best agreement.

Figure 4B shows the correspondence index (CI) across all modalities for the NKI dataset, which measures the degree to which the evidence agrees across all modalities. Overall, positive CI (how well modalities agree that an edge exists) was higher than negative CI (how well they agree an edge does not exist), but cross-modal disagreement predominated (mean inverse $CI = 0.63$, mean positive $CI = 0.25$, mean negative $CI = 0.11$). Positive CI was highest for short connections (0.49), while negative CI was highest for long connections (0.25). Ipsilaterally, CI was mainly positive, while for heterotopic contralateral connections positive and negative CI were roughly equal. For homotopic connections, mean positive CI was 0.65, with zero negative CI (indicating perfect agreement on the existence of a connection for 65% of homotopic pairs). This measure is shown geometrically in Figure 4C, both as a superior aspect 3D rendering and as horizontal 2D sections. Qualitatively, the pattern of positive CI resembled a largely symmetric lattice, including edges between locally proximal regions, as well as most homotopic regions. Negative CI was largely confined to long-range, medial, anteroposterior edges. Inverse CI (disagreement) had no distinct regularity or geometric pattern; edges were generally long-range and diffuse throughout the geometry of the brain, and included both contralateral and ipsilateral connections. Figure 4D shows the mean CI per ROI. Lateral temporal lobe and the temporoparietal junction had the highest positive CI, with CCa, CCp, and dorsolateral PFC having moderately high values. HP and CCs had the highest negative CI, while medial PFC, medial temporal lobe, V1, and Ia had the lowest correspondence. These distributions were very similar for the ICBM dataset (Figure S1).

It is instructive to examine the CI in specific components of the connectome, for instance by looking only at the DMN, a set of regions which are most likely to be engaged in the resting state (Greicius et al., 2009). Notably, for the NKI dataset, none of the human modalities matched the structure of the CoCoMac-derived DMN well (Figure 5). All three imaging networks failed to find connections between PFCm and PHC, TCv, and CCr, which were dense connections in CoCoMac. The fMRI network was comprised mainly of strong homotopic connections, including one between the left and right CCr, which does not exist according to CoCoMac. The DWI network lacked a core set of contralateral connections involving PHC and TCv. Like fMRI, the cortical thickness network had edges between left and right CCr, and like DWI, it lacked most connectivity for HC, PHC, and TCv. These differences are summarized by the CI, as shown in Figure 5 (top left). Figure S2 shows an alternative visualization within the DMN; specifically, connectivity estimates for each modality rendered on the ICBM surface for the left PCm.

Figure 6B shows normalized ACC, across network densities, for all human modalities, both parcellation schemes, and fMRI with and without GSR. ACC followed the same pattern as correlation (Figure 6A), and this was generally consistent across densities. Interestingly, R_{fmri} and $R_{\text{fmri}-\text{gsr}}$ had similar agreement with both R_{dwi} and $R_{\text{thickness}}$ at low density ($\sim$ 23%), but this became increasingly worse at higher densities. Also following the correlative results, ACC was consistently worse for estimates produced using the Cambridge parcellation than for RM, and CI was also generally worse for Cambridge (albeit moderately) as well (Figure 6C). Finally, ACC between R_{fmri} and $R_{\text{fmri}-\text{gsr}}$ was comparably high for both parcellations (0.6 to 0.7), which also concords with the correlative findings. As for the weighted comparisons, $R_{\text{thickness}}$ had a remarkably poor within-modality reliability for split-half binarized comparisons, with an ACC of approximately 0.6 at low density, reducing to 0.43 at the maximal CoCoMac density of 58% (Figure S8).

3.4. Clustering

K-means and hierarchical clustering allow us to compare the ways in which networks form stable clusters on the basis of their edge weights. Figure 7A shows the stability matrices for k-means clustering of each modality, which quantifies the probability, based on 500 subsamples, of any two ROIs being clustered together at $k = 12$ partitions. The consensus matrix is also shown (far right), which represents the likelihood of two ROIs being clustered together across modalities. The matrices are shown both in their original ordering (first row), and ordered according to the consensus clustering solution for $k = 12$ partitions (second row). Figure 7B shows the 12 partitions derived from

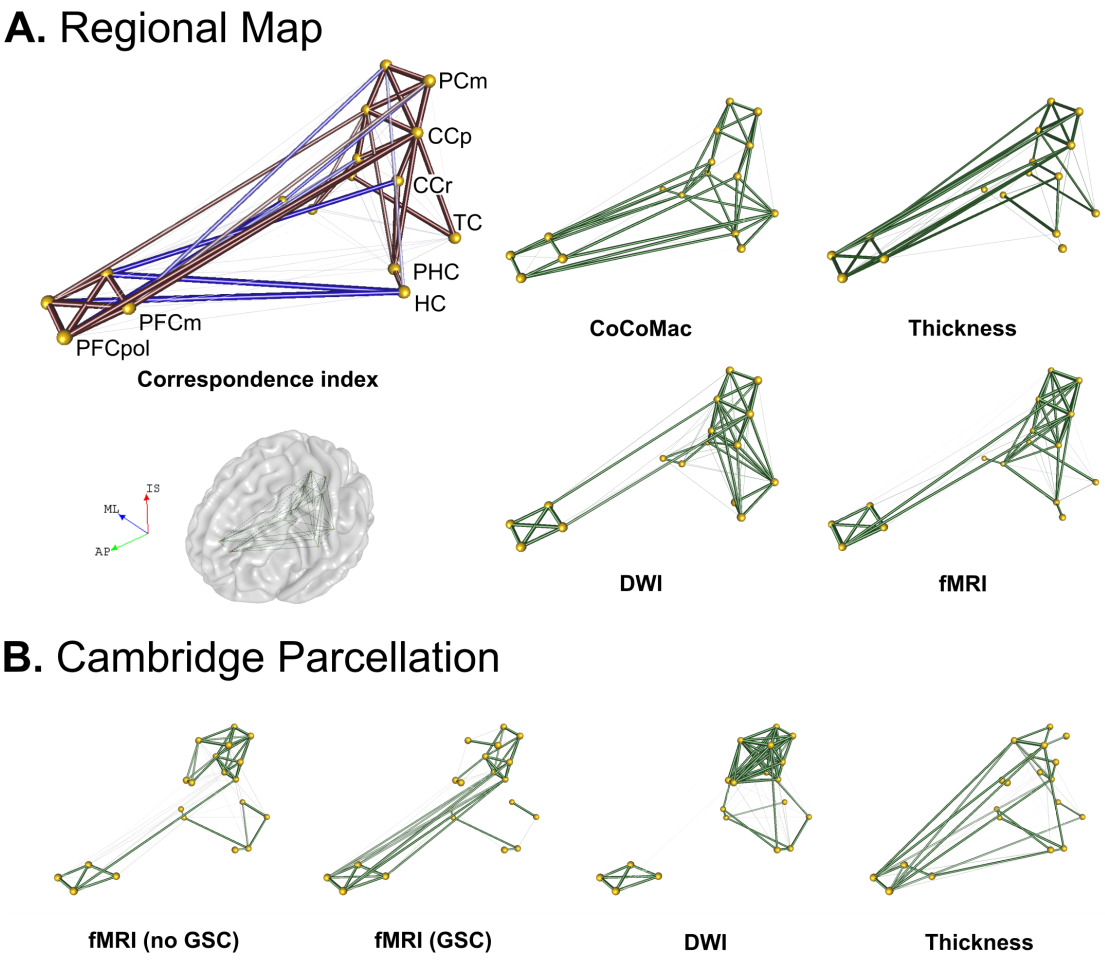


Figure 5: **The default mode network (DMN). A.** At right, networks for the RM parcellation are shown for each of the four modalities. Darker, larger, and more opaque edges signify stronger edge weights. At top left, the CI is shown, as a measure of correspondence across these four modalities. Red edges signify stronger positive CI, while blue edges signify stronger negative CI; smaller, more transparent edges are ones with more disagreement. ROI labels are given in Table 1. The cortical surface at bottom left is shown for orientation. **B.** Renderings of the DMN network for the human modalities using comparable regions of the Cambridge parcellation. At left, fMRI networks derived with and without a GSR step are shown.

the consensus clustering solution, painted on the ICBM template cortical surface. These images allow a number of qualitative observations. Homotopic connections were among the most stable, which likely accounts for the interhemispheric symmetry of the partitions. Ipsilateral and contralateral clustering patterns were nearly identical to one another for both R_{fmri} and $R_{\text{thickness}}$, while clustering was much more stable ipsilaterally than contralaterally for R_{cocomac}, and R_{dwi} was intermediate in this respect. As shown in Figure 7C, R_{fmri}, R_{dwi}, and R_{cocomac} had comparable stability across all partition sizes (measured by mean stability contrast, σ, which measures the ratio of within-cluster to between-cluster connectivity), whereas $R_{\text{thickness}}$ was substantially less stable across subsamples.

For the consensus clustering solution at $k = 12$, the partitions were mostly spatially contiguous, and apart from the dorsolateral PFC (orange and turquoise) and occipital cortex (dark blue and yellow) all were symmetric across

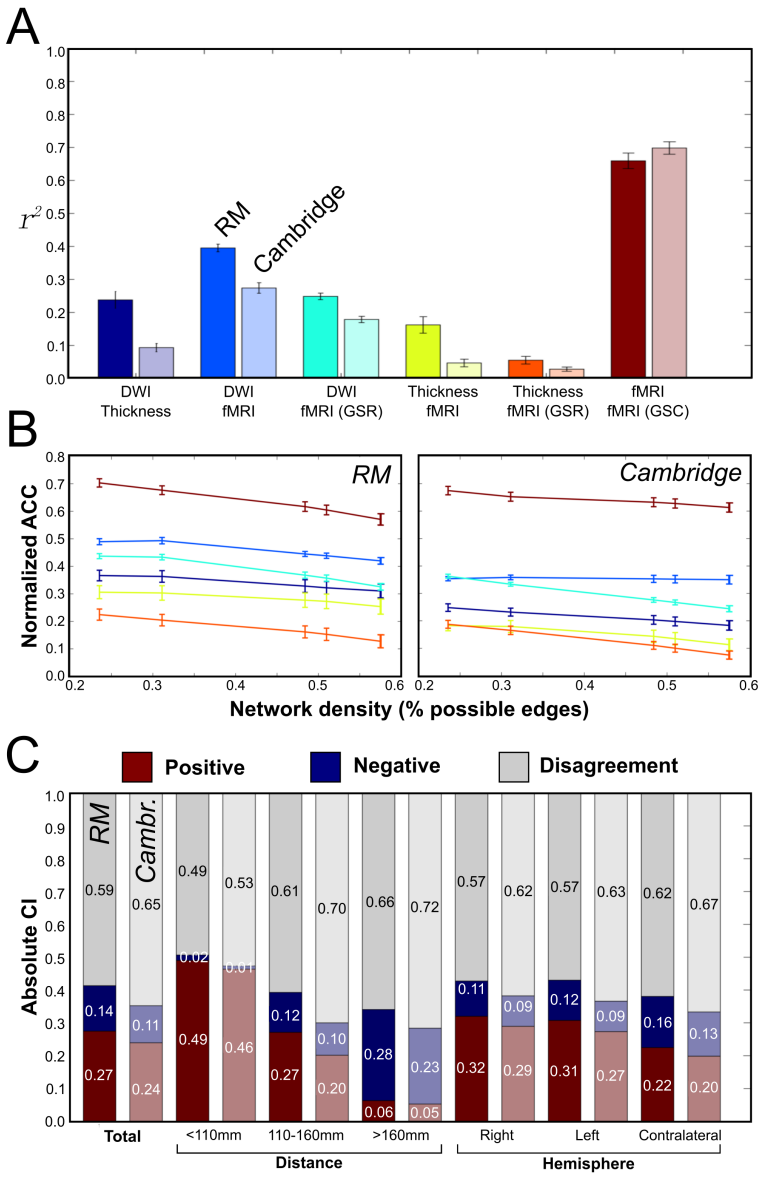


Figure 6: **Summary of comparisons for human modalities only. A.** Weighted (correlative) comparisons of connectivity estimates for the Regional Map (RM) and Cambridge parcellations (lighter bars, as labelled), where GSR refers to global signal regression corrected fMRI. **B.** Mean (over subsamples), normalized ACC shown across all CoCoMac densities, for RM and Cambridge parcellations. The colour coding is the same as for (A). Errors bars show the standard deviation across subsamples. **C.** Mean CI for both parcellations (as labelled), derived at the maximal CoCoMac density of 58%. Means are displayed for all edges (total), and also divided by distance and hemisphere.

hemispheres. The only noncontiguous partition (lighter orange) was comprised of posterior insula, temporal pole, and medial temporal lobe. These ROIs also had very low mean stability contrast for the consensus clustering (Figure 7D), along with orbitofrontal cortex, cingulate cortex, precuneus, lateral temporal lobe, and the temporoparietal junction. Figure S4 shows the partitioning solutions for each individual modality, which gives more insight into

their correspondence. In general, some of the poor consensus can be attributed to the separation of clusters by hemisphere for R_{dwi} and R_{cocomac}, but not for R_{fmri} and $R_{\text{thickness}}$. However, a large degree of divergence was also observed ipsilaterally, and all four modalities give rise to distinct solutions. The clustering pattern at $k = 12$ can be further elucidated by considering different levels of partitioning sizes. Figure S6 shows solutions for $k = 8$, 12, and 16. Most of the divergent, modality-specific clustering patterns were already present at $k = 8$, and these patterns tended to further subdivide with increasing k. Generally, this subdivision was symmetric across hemispheres, but in some cases – particularly for R_{fmri} – the new partitions were asymmetric.

Figure S5 shows the same results for the hierarchical clustering approach. The consensus solutions for both methods had substantial overlap (compare Figure 7B and S5B), but also exhibited a number of distinct differences. For instance, the stability of R_{cocomac} was very poor under hierarchical clustering, while R_{fmri} was greatly improved, and $R_{\text{thickness}}$ showed an improvement with increased values of k. Only R_{dwi} appeared similarly stable under both approaches. Interestingly, stability contrast for the consensus clustering was substantially higher for hierarchical versus k-means, specifically in precuneus, posterior cingulate, and primary somatosensory cortex (Figure 7D and S5D), while the k-means consensus was better for medial and left lateral prefrontal cortex.

3.5. Signal-to-noise

Signal-to-noise ratio SNR for the three human modalities is shown in Figure 8, for both ICBM and NKI datasets, along with data completeness DC for CoCoMac. Notably, SNR was higher for NKI than for ICBM, for all three modalities, providing a quantitative indication of the improved data quality of the NKI dataset. Nonetheless, the overall pattern of SNR was similar across datasets for the same modalities, and differed substantially between modalities. The relationship of normalized SNR and DC to the correspondence index CI was analyzed by the Spearman rank coefficient ρ. SNR for all three human modalities were significantly correlated with CI, while DC for CoCoMac showed no relationship (ρ = 0.29,0.26, 0.32, and 0.01 for DWI, fMRI, Thickness, and CoCoMac, respectively). Figure 8 also shows the normalized SNR, averaged for each ROI, as a distribution on the ICBM template surface. This distribution bears a remarkable qualitative resemblance to the distribution of mean CI (Figure 4D).

4. Discussion

The assignment of weights to edges is a critical step in inferring connectivity from empirical evidence. When constructing a graph by assigning weights to edges, and treating this graph as a representation of a physical network which transmits information, we are implicitly assuming that edge weight informs us about the efficacy of information transfer. This assumption, however, is rarely tested. Here, we compared cortical connectomes from four different modalities, using both weighted and binarized measures of edge strength. As a general finding, and given the assumption that all connectivity methods are attempting to model the same underlying physical connectivity, correspondence between all four modalities was poor. This highlights the difficulty of drawing conclusions about the *physical connectivity* of the human brain based upon any one neuroimaging modality. On the other hand, if we consider the degree of difference between the aspects of physical connectivity modelled by each modality, we do find a relatively impressive correspondence, in particular between the DWI (structural) and fMRI (functional) approaches, which achieve approximately 40% shared variance. When we consider binarized networks, CoCoMac and DWI also show a relatively good correspondence. Moreover, the use of higher resolution data (NKI) appears to substantially increase the degree of correspondence between all three human imaging (but not CoCoMac) modalities. Because we have both species and modality differences, along with a variety of methodological issues associated with each modality, the task of disentangling these results is a difficult one. Nonetheless, the present comparisons have a number of important features which can be used to better characterize what inferences can – and cannot – be drawn from the evidence of each modality. In what follows, we discuss the specific results of weighted and binarized

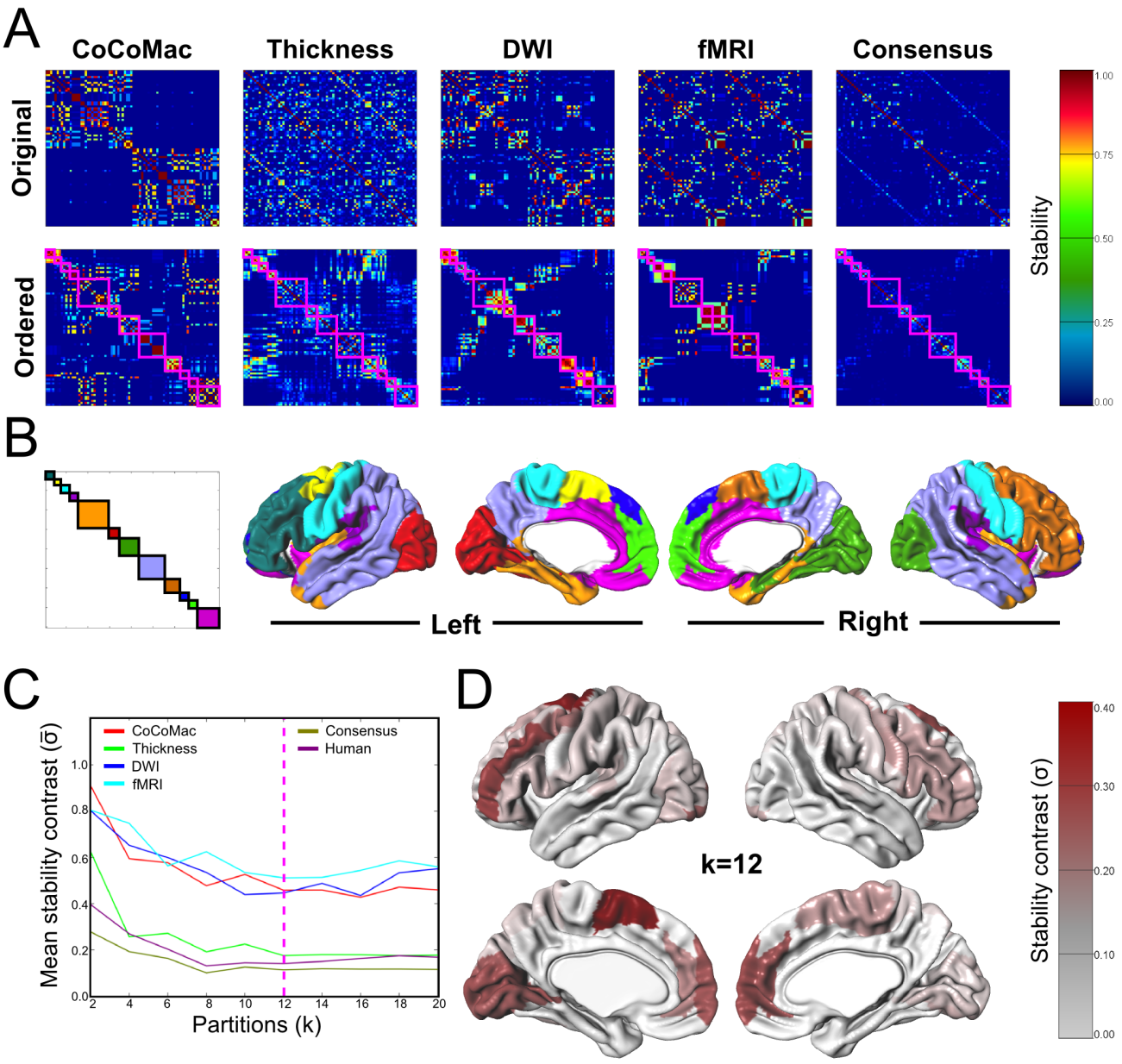


Figure 7: **K-means clustering results at $k = 12$ partitions. A.** Stability matrices for each modality (how frequently ROI pairs cluster together, within modalities) and the consensus matrix (far right; how frequently ROI pairs cluster together, between modalities). Matrices are shown in their original order (top), and ordered by the consensus clustering (bottom), where the 12 clusters are outlined in pink. **B.** Consensus clusters at $k = 12$, painted on the ICBM template cortical surface. **C.** The mean stability contrast (σ) for each modality, at different values of k. The dashed pink line shows $k = 12$, which was used here. **D.** ROI-wise σ, painted on the ICBM template cortical surface. Higher σ (red) indicates an ROI is more strongly clustered within its assigned cluster than between other clusters.

comparisons and hierarchical clustering, and finally consider the strengths and limitations of each modality, which may account for their agreement, or lack thereof.

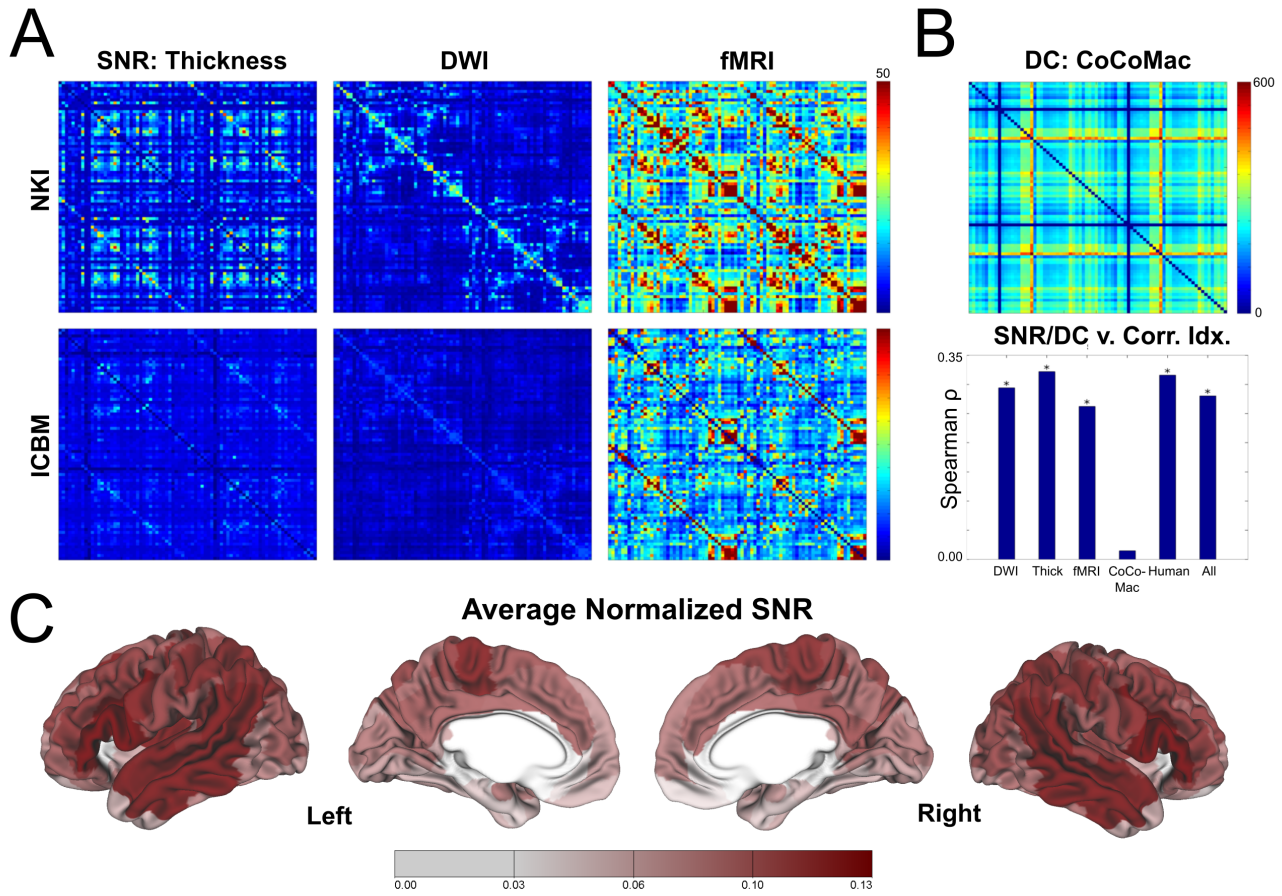


Figure 8: **Signal-to-noise ratio (SNR) and data completeness (DC). A.** Top row shows SNR averaged over ROI pairs, for each human imaging modality derived from the enhanced NKI dataset. Second row shows the same for the ICBM dataset; notably, all values are smaller for ICBM, but patterns are constant across datasets for each modality. **B.** At top, DC is shown for the CoCoMac network; i.e., the average number of injections and labelled sites for each ROI pair. At bottom is the Spearman rank coefficients for SNR/DC versus the correspondence index CI, obtained across all modalities. Notably, significant correlations were found for the human imaging modalities, but not for CoCoMac. **C.** Average SNR, normalized for each modality, plotted on the ICBM template surface. The distribution bears a resemblance to the distribution of mean CI, shown in Figure 4C.

4.1. Correspondence across modalities

For both weighted and binarized comparisons, we find more divergence than convergence between the modalities, with $R_{\text{fmri}}/R_{\text{dwi}}$ having the best agreement in both cases. This degree of correspondence (i.e., between DWI and fMRI) is comparable to that reported by earlier intermodal comparisons (Honey et al., 2009). Interestingly, in studies which compared functional and structural connectivity for generative models (i.e., where the underlying structure is known *a priori*), average correlations in the range of 0.6-0.7 (Honey et al., 2009; Messé et al., 2015) were achieved. This indicates that the empirical correspondence observed here is likely about as high as can be expected. For both comparative approaches, correspondence was highest for shorter, ipsilateral, and homotopic contralateral connections (Figure 3B and Figure 4B), with the latter having a positive CI of 0.65, indicating cross-modal agreement for 65% of these edges. These findings largely agree with those of Honey et al. (2009), who compared DWI-based SC, and fMRI iFC. Positive CI was distributed in a regular, lattice-like set of edges, along with most

homotopic contralateral edges (Figure 4C), indicating that local and homotopic connections have the best agreement across all modalities. Negative CI was distributed mainly along longer, medial edges, both ipsi- and contralaterally, while disagreement appeared to be highest for long heterotopic contralateral and anteroposterior ipsilateral edges (Figure 4C). Interestingly, the distribution of mean positive CI across ROIs was highest in heteromodal regions (Figure 4D), while S1, V1, medial temporal, and medial frontal regions had relatively poor correspondence. The distribution of positive CI bears a striking resemblance to that of individual variability in fMRI iFC, as well as the degree of evolutionary cortical surface expansion between the macaque and human cortex (Figure 3 of Mueller et al., 2013). The reasons for this resemblance are not immediately obvious, since one might expect regions with a greater degree of evolutionary divergence to have a worse correspondence, particularly with CoCoMac. However, our evidence suggests that local ipsilateral connectivity and homotopic contralateral connectivity (i.e., those most likely to form network modules) are largely preserved across species, and stable across modalities, while longer, heterotopic contralateral and anteroposterior connections (i.e., those most likely to act as intermodule connections) have the worst stability. This pattern suggests that regions of cortex that are less locally interconnected, and/or have more long-range connections, are less likely to have good correspondence across species and modalities.

It is noteworthy that the binarized comparison improves considerably (approximately 10%) when considering human modalities only (Figure 6). This is not very surprising, given the very different nature of the CoCoMac data, the fact that it is derived from a very different primate species, and the observation that CoCoMac compares poorly with the two correlational modalities (see Figure 4). It is unclear, however, whether this improvement in correspondence is attributable to neuroimaging methods converging on a "true" human connectome, or simply that they are affected by similar biases. The correlational modalities, for instance, are likely to both capture brain symmetry for different reasons (one functional and one morphological). The propensity to find short connections may also result in improved correspondence, again for different reasons (a distance bias in DWI, and smoothing effects for fMRI and cortical thickness). However, it is likely that much of the improvement reflects a real convergence on the same underlying connectivity pattern, which is reduced by comparison to CoCoMac both because of actual species differences and the difficulty of matching homologous regions (discussed in more detail below). Future studies which focus more directly on comparisons between estimates of human connectivity will likely better elucidate these issues.

4.2. Correspondence across datasets

At the time of writing, the MRI techniques utilized to obtain the NKI dataset are state-of-the-art, including a 3T scanner and novel acquisition sequences that produce high temporal and spatial resolution fMRI, high angular resolution DWI, and good T1 contrast for cortical thickness analysis. In contrast, the ICBM dataset is rather outdated, using a 1.5T scanner and yielding - by today's standards - images with low spatial, temporal, and angular resolution. Contrasting these two datasets allows us to assess the impact of improved MRI techniques on the intermodal correspondence of connectivity analyses. Overall, the pattern of both correlative and binary comparisons was consistent across datasets, indicating a stability to the relative degree of intermodal correspondence. We also observed a marked increase for NKI in both intermodal correlations and binary correspondence for the human imaging modalities, but not for comparisons with R_{cocomac} (Figures 3B and S5). The reasons for this poorer correspondence are not clear; speculatively, it may be that the NKI dataset better captures real interspecies divergence. There are, on the other hand, several possible explanations for the observed improvements. Notably, increased angular resolution for DWI allows a more well-defined orientation distribution function, which can improve the ability of tractographic algorithms to find pathways through highly isotropic tissue (i.e., containing crossing, kissing, or fanning fibres). Increasing temporal resolution for rs-fMRI allows a more temporally precise analysis of regional coupling strength, which improves the ratio of signal-to-noise in correlational analyses. Additionally, increased field strength allows a better T1-weighted contrast, which improves the ability of cortical thickness algorithms to perform tissue classification and detect cortical boundaries. These results argue in favour of the benefit of applying state-of-the-art imaging techniques to datasets intended for use in analyzing connectivity using one or more modalities. However, they also demonstrate that, despite these improvements, there remains a good

deal of disagreement between connectomes derived from these modalities, which is elaborated upon below.

4.3. Signal-to-noise

The signal-to-noise ratio (SNR) reflects the degree to which a salient signal can be isolated from the noise present in an observed distribution. In the present study, we evaluated this quantity as the degree to which connectivity estimates obtained from subsamples exceeded variance across them. This allowed us to evaluate how well cross-modal correspondence could be predicted by the stability of the estimates. For CoCoMac, SNR was modelled as the number of studies contributing to information about a particular connection, referred to as data completeness (DC). In contrast to the human modalities, DC did not significantly predict CI (Figure 8B). We found that the SNR of each of the human imaging modalities predicted the CI, with Pearson coefficients between 0.3 and 0.4. This relationship suggests that the stronger and more stable a correlative or structural relationship, the more likely it is to correspond across functional and structural modalities. This finding is in accordance with a similar observation by Honey et al. (2009), that edges with structural connectivity, determined by DWI tractography, had less RS-FC variability both across and within sessions.

Interestingly, fMRI exhibited a much higher SNR than either DWI or thickness. This highlights the degree to which these modalities differ in their connectivity estimates. fMRI, based on BOLD time series over time, gives rise to relatively stable patterns of covariance, whereas cortical thickness, which captures morphological patterns over much longer time scales, and is likely influenced by multiple factors, including genetics and life experience. Cortical thickness is also a noisy measure, and results from various approaches have been shown to have substantial disagreement (Martinez et al., 2015). Probabilistic DWI tractography, on the other hand, attempts to sample a complex diffusion space, and even at the dense sampling rate utilized here, appears to have a much higher degree of variance than fMRI activity. The spatial distribution of the SNR effect, obtained by averaging coefficients for each region, is also shown as a surface rendering in Figure 8C. Notably, this pattern shows a strong similarity with the surface rendering of mean CI plotted in Figure 4D, indicating that the correlation between SNR and cross-modal correspondence is also reflected in a similar spatial distribution. Finally, when comparing between the NKI and ICBM datasets, we found a significantly higher SNR for all three modalities (Figure 8A) for NKI. This observation reinforces the idea that the improved cross-modal correspondence we observe for the NKI dataset is attributable to an improved signal quality in this higher-quality data set.

4.4. K-means and hierarchical clustering

Both k-means and hierarchical clustering approaches allow a network graph to be partitioned into distinct subsets, based upon some distance measure. In the present case, we considered edge weights to confer similarity (the inverse of distance), and computed clustering solutions for each modality, across 500 subsamples, to determine the reliability with which each pair of ROIs was clustered together. We further computed a "consensus" clustering, which determined how reliably ROI pairs appeared in the same cluster across modalities, thus arriving at a set of partitions derived from the integrated evidence of all four modalities. For k-means, we found that R_{fmri}, R_{dwi}, and R_{cocomac} had similar levels of stability across subsamples ($\sim 0.5 - 0.6$ for $k > 6$), while $R_{\text{thickness}}$ had relatively poor stability (~ 0.2). In contrast, hierarchical clustering resulted in disparate levels of stability for each modality, with the general trend $R_{\text{fmri}} > R_{\text{dwi}} > R_{\text{thickness}} > R_{\text{cocomac}}$. R_{cocomac} in particular performed very poorly using hierarchical clustering, which suggests that this approach is not ideal for such a comparison. The discrepancy between methods highlights the degree to which a solution is dependent upon the choice of algorithm. Notably, such algorithmic dependency has been previously characterized for fMRI data specifically, for both a coarse parcellation (Marrelec et al., 2015) and a voxel-wise approach (Thirion et al., 2014). In both studies, hierarchical clustering performed well, while k-means was less accurate or unstable; this was also likely the case here. For R_{cocomac} in the present study, the discrepancy between hierarchical and k-means solutions is possibly due to the particular means by which distance measures are used to form clusters in the two approaches. Hierarchical clustering attempts to make

binary decisions about pairs of regions, based upon their distance measure (in this case, using the Ward criterion). Because R_{cocomac} is a discrete matrix (i.e., comprised of integer values), and highly skewed towards maximal values, this approach results in a high number of arbitrary decisions made on pairs with identical distances, which must be resolved in a pseudorandom way. As a result, the multiple runs on R_{cocomac} produce highly varying results. This ambiguity is likely avoided to some extent by the k-means approach, in which regions are added to a cluster on the basis of their distance from a centroid for that cluster. In other words, this approach attempts to assign borders which maximize differences between sets of regions, which reduces the number of arbitrary decisions that must be made about individual pairs of regions. To address this issue, future studies might evaluate alternative formulations of edge weighting for the CoCoMac database (see CoCoMac network, below).

The consensus clustering solution for k-means, shown in Figure 7B, provides a set of spatially contiguous, hemispherically symmetric clusters. Figure 7D shows the stability of this clustering solution across subsamples. Together, these results reinforce the higher correspondence of local connectivity, and also highlights regions of cortex where modalities disagree most strongly; specifically, CCa, CCp, orbitofrontal cortex, the temporal lobe, and the temporoparietal junction. This pattern contrasts somewhat with that of mean positive CI (Figure 4D), which reflects the very different aspects of connectivity assessed by these methodologies. In particular, CI is a binarized comparison, while hierarchical clustering uses edge weights as continuous values to determine how ROIs cluster together. The disagreement of lateral temporal lobe, for instance, is due to its being separated by hemisphere for R_{cocomac} and R_{dwi}, but not for R_{fmri} and $R_{\text{thickness}}$. This discrepancy is likely due to homotopic connectivity; the latter modalities produce high weights for homotopic connections, which drive their clustering behaviour but are reduced to a binary score in the CI calculation, equivalent to weaker but more prevalent local edges. The poor clustering of medial temporal lobe, on the other hand, agrees with the mean CI result, and suggests that this part of cortex has the least consistent connectivity structure. Cross-modal comparisons using clustering methods have been previously reported for limited regions of cortex. Kelly et al. (2012b), for instance, focusing on the human insula, demonstrated a convergence of clustering solutions across partitioning sizes, for both resting state and task-evoked fMRI, as well as covariance in GM density. In the present study, we do find a qualitative similarity between R_{fmri} and $R_{\text{thickness}}$ solutions (see Figure S6).

It is interesting to examine how partitioning compares across different choices of k. From Figure 7C it is clear that there is no peak, or optimal k at which the data can be clustered; in other words, past $k = 8$, further subdivision of partitions neither adds nor subtracts from the stability of the solution. To better assess how k affects the partitioning solution, Figure S6 illustrates solutions for $k = 8$, 12, and 16. Notably, the idiosyncratic clustering behaviour of each modality is largely already apparently at the $k = 8$ solution. For instance, the splitting of medial temporal lobe and visual cortex is apparent for R_{cocomac} at this level, rather than occurring at a higher parcellation, which would indicate it arises from forcing a subdivision in a way unnatural to the data. On the other hand, for most modalities clear instances of splits which impose hemispheric asymmetry, or "unnatural" seeming parcels can be found, mostly in regions exhibiting poor mean stability contrast (Figure 7C). This highlights a limitation of the present approach. The use of a coarse parcellation scheme such as RM to perform such clustering, while providing an informative perspective on the relationship between clustering behaviour for different modalities, would likely benefit from a complementary analysis with a focus on voxel- or vertex-wise clustering. Future studies could also provide a more comprehensive characterization of the behaviour of various clustering approaches, parcellation schemes, and cluster sizes in assessing multimodal connectivity.

4.5. CoCoMac network

The CoCoMac database provides us with a very useful set of metadata which facilitates comparisons between histologically confirmed anatomical connectivity of rhesus macaques, and methods which indirectly predict connectivity of the human brain. This potential is limited, however, by a number of factors. Notably, the use of parcellation schemes with different naming schemes and spatial scales results in the difficult issue of how individual schemes should be mapped together (Bakker et al., 2012). While this issue has been addressed with textual

(Stephan et al., 2000) and spatial (Bezgin et al., 2008) approaches, it remains a significant source of uncertainty. Another issue is the number of ROI pairs for which no experimental data has been entered into the database; these consist primarily of contralateral pairs, 58.7% of which are missing information. Where possible (e.g., correlative comparisons and the computation of CI), we treated these edges as missing data, but for thresholding purposes unknown edges were treated as nonexistent, which is undoubtedly a factor in its much smaller ratio of contralateral-to-ipsilateral connections, compared to the other modalities (Figure 2F). The nature of this sampling bias could be very informative for future applications of CoCoMac data. Notably, due to the sheer size of literature relating to tract tracing experiments, and its continued growth, not all the evidence has yet been entered into the database. Another explanation is that a large majority of tracing studies – particularly early ones – focused exclusively on ipsilateral connectivity, often even neglecting to specify the hemisphere in which they were performed. At present, of 43,746 labelled sites entered in CoCoMac, only 2,691 (5.8%) are contralateral (personal correspondence). A third possibility is that many contralateral connections were not investigated because there was simply no evidence to suggest their existence, and thus motivate an experiment. However, while signal-to-noise in the human imaging modalities was related to cross-modal correspondence, the quantity (or completeness) of data in CoCoMac was not found to have such a relationship (see Supplemental Figure 5). Thus, where information does exist for CoCoMac, its quantity does not appear to have an effect upon the degree of its correspondence with human imaging data.

A second limitation of CoCoMac is the fact that it is based on rhesus macaques, and thus for any comparison with human connectivity it is difficult to differentiate species and modality differences. This question of homology is a well-known issue and has been the subject of numerous lines of research (reviewed in Passingham, 2009; Petrides et al., 2012). Human and macaque surfaces were coregistered by identifying homologous morphological landmarks and warping between them (Van Essen and Dierker, 2007). In this process, heteromodal regions (particularly the temporoparietal junction and dorsolateral prefrontal cortex) were expanded to a greater degree than unimodal regions, raising the possibility that these regions evolved novel connectivity patterns to subserve the unique cognitive functions of humans. To investigate this question, Mantini et al. (2013) compared BOLD iFC in rhesus macaques and humans, finding a topological correspondence between sensorimotor and associative networks, but a discrepancy in the functional activation of these networks, particularly in associative regions. This evidence suggests that functional networks are preserved between macaques and humans, while the patterns with which they are activated can differ. The authors additionally found two lateralized frontoparietal networks in human subjects, which had no equivalent in the macaques, indicating some evolutionary divergence in network topologies, as well. This divergence may account in part for the disagreement in long, ipsilateral anterioposterior edges observed in the present study. In a separate approach, Margulies et al. (2009), seeding in different parts of the human precuneus, found a correspondence between BOLD iFC and the distribution of injected or labelled sites in a number of independent macaque tracing studies. Similarly, Kelly et al. (2010), using RS-fMRI data-driven clustering to define their ROIs in human subjects, found a qualitative correspondence between iFC of humans and previously acquired macaque tract tracing findings (Petrides and Pandya, 2009), such that iFC strength corresponded to the spatial extent of labelled cells. These results suggest that, at least for specific ROIs, macaque and human connectivity are comparable. However, while in the present study both clustering results (Figure S6) and precuneus connectivity (Figure S5) for R_{cocomac} and R_{fmri} do have a qualitative correspondence across much of cortex, we found a generally poor quantitative agreement between them. Interestingly, we also found that correspondence is not uniform across the cortex. Indeed, the precuneus, prefrontal cortex, and lateral temporal lobe, investigated by Margulies et al. (2009) and Kelly et al. (2010), are amongst ROIs with the best positive correspondence between R_{cocomac} and the other modalities, including R_{fmri}.

For CoCoMac, we have the most direct evidence of physical connection strength in the form of relative label density (see the CoCoMac Data Entry Manual, http://www.cocomac.org). However, the assignment of label density is based upon the translation from a qualitative (or loosely quantitative) assessment obtained from individual articles, to an integer range of 0 to 3. Since each study had its own variation of scope, parcellation schemes, and experimental protocols, the degree to which density values correspond across literature entries is highly uncertain. For instance, a study with a very limited scope is likely to have been formulated to maximize label density in the target ROI, while for an experiment of larger scope, this density might be much sparser relative to other ROIs. Since, in this study, a substantial majority of CoCoMac connections have been assigned maximal densities, this issue

highlights an important limitation of our approach. Additionally, as many articles do not provide any information about label density, a further assumption must be made regarding the assignment of edge weights for connections which are present, but whose density is unknown. We decided to assign assume an "expected value" of 2 for these edges, but given the skewed distribution of CoCoMac, a value of 3 might have also been assigned, which would undoubtedly have altered the resulting network. Recently, Markov et al. (2012, 2013) have helped address this issue by performing systematic tracing experiments in which label density was quantified as the fraction of all labelled neurons (FLN) in a given experiment. Considering only ipsilateral ROIs, these authors also report a high connection density (66%), and a lognormal (i.e., skewed) distribution of FLN values, which corresponds here to the distributions of both R_{cocomac} and the untransformed R_{dwi}. Such data has a promising role as a more meaningful weighting scheme constraining anatomical macaque networks. Indeed, a recent study (Miranda-Dominguez et al., 2014) compared human and macaque connectivity using FLN data from the Markov et al. study. The authors report a better correspondence between macaque anatomical and human RS-fMRI connectivity than is found in the current study ($r^2 = 0.417$ for whole brain ipsilateral connections, versus $r^2 = X$ here). Correspondence improved slightly ($r^2 = 0.293$) when only the top 15% of connections were considered. These findings suggest that such refined tracing measurements can indeed improve inter-species correspondence, although it remains to be seen how the inclusion of contralateral FLN values might affect this picture.

4.6. Cortical thickness network

Correlations in cortical thickness have potential for inferring connectivity at the group level, by quantifying the degree to which neural morphology (i.e., GM growth and atrophy) is coordinated across the cortex. Numerous lines of evidence support the idea that GM morphology is affected by activity-dependent processes (reviewed in Draganski and May, 2008). Accordingly, brain regions which are strongly connected are likely to have more similar activation patterns, and corresponding patterns of GM morphology. Due to this, correlations in cortical thickness might be expected to resemble FC. Because both are correlational, they may also be influenced by indirect connections and common driving inputs (i.e., parallel processing streams). $R_{\mathrm{thickness}}$ was indeed most closely related to R_{fmri}, in terms of its contralateral connectivity, strong homotopic connections, edge weight distribution (Figure 2), and clustering behaviour (Figure S6). However, in general $R_{\mathrm{thickness}}$ had a relatively poor quantitative correspondence (either correlational or binary) with any of the other modalities. This is mainly attributable to longer and contralateral edges, while its best correspondence occurred for short connections – especially local ipsilateral connections.

The poor correspondence of $R_{\mathrm{thickness}}$ may result from a number of limitations for this approach. Notably, when evaluated within-modality using split-half analysis, $R_{\mathrm{thickness}}$ showed a poor reliability for both weighted and binarized comparisons (Figure S8), which indicates that there is a good deal of random variation in the correlative structure of individual samples themselves, and can partially account for the lack of cross-modal correspondence. The estimation of cortical thickness is a complex problem, based upon a series of processing steps which each come with their own set of assumptions relating the T1-weighted intensity profile to the underlying tissue (Sled et al., 1998; MacDonald et al., 2000; Kim et al., 2005). It is therefore important to consider how these various assumptions might influence our results. In particular, the use of global intensities to segment images into tissue classes and determine the final position of cortical boundary surfaces depends on an assumption that the composition of each tissue class is uniform across the cortex, which is unlikely to be true. Tissue contrast will be poorer, for instance, where GM and WM are mixed, such as in M1, which contains a disproportionately high amount of myelinated fibres (Geyer et al., 2011). Similarly, cortical regions with more convoluted (within subject) or variable (between subjects) morphology are less likely to support a good approximation or between-subjects correspondence, respectively. All of these factors can potentially affect correlations; interestingly, regions for which $R_{\mathrm{thickness}}$ has poor correspondence include M1, medial prefrontal cortex (which has a high degree of gyrification and individual variation; Zilles et al., 1988), medial temporal gyrus and insula (which have poor tissue contrast), and medial occipital lobe (whose highly convoluted geometry is difficult to model). This possibility suggests that improvements to cortical thickness estimation software may be reflected in a better correspondence between $R_{\mathrm{thickness}}$ and the other modalities.

The inference connecting morphological correlations to connectivity is also confounded by other possible sources of correlation. For instance, GM morphology is known to have a strong genetic influence (Baaré et al., 2001; Thompson et al., 2001; Wright et al., 2002), which is independent of environmental influence and is likely to produce covariance which cannot be explained by the activity-dependence hypothesis. Additionally, correlations due to indirect influences, such as mutual source ROIs or parallel contralateral signalling pathways, are also likely to produce correlations in the absence of direct anatomical connections. Because the brain is generally symmetric across hemispheres, we can expect a high degree of correlation in morphology between homotopic regions, regardless of whether these are directly connected or not. Indeed, the lower degree of interhemispheric connectivity found in either R_{cocomac} or R_{dwi} may account for many of the false positives which contribute to lower weighted or binary comparisons here, as well as decreasing the clustering consensus. While statistical methods exist to address this issue, such as partial correlation (Smith et al., 2011), these typically have the disadvantage of being too conservative to discover a whole-brain connectome which is as densely connected as is implied by the macaque tracing evidence. This further supports the idea that modalities must be combined to properly exploit their complementarity and isolate those edges for whose existence multiple lines of evidence agree.

4.7. DWI network

DWI probabilistic tractography is a powerful technique for indirectly observing anatomical connectivity in the human brain. In this study we used a computationally intensive approach (seeding 10,000 times in each WM voxel) in order to address the bias imposed by long distances and regions of uncertainty (i.e., crossing, kissing, or branching fibres). As a result, we obtained a connectivity matrix of higher density than has typically been reported for similar studies (e.g., Honey et al., 2009; Gong et al., 2012). R_{dwi} had the best correlative and binary correspondence to R_{fmri}, sharing 41% of its variance. Qualitatively, R_{dwi} is highly symmetric across hemispheres, with a higher proportion of ipsilateral than contralateral connectivity, and high distance-weighted betweenness centrality in cingulate regions, similar to that of R_{cocomac}. For both R_{fmri} and $R_{\text{thickness}}$, which have higher contralateral connectivity, correspondence was much better for short, ipsilateral connections. Gong et al. (2012), who also compared DWI and cortical thickness networks, report a similar convergence for ipsilateral, and divergence for contralateral connections. While the present approach cannot completely rule out the influence of methodological distance bias on this pattern, it is tempting to speculate that the imbalance between ispi- and contralateral connections is real, and that the even ratio found for R_{fmri} and $R_{\text{thickness}}$ is a result of parallel processing streams, subcortical modulation, and (in the case of $R_{\text{thickness}}$) common genetic influences. Markov et al. (2012, 2013) observe that connectivity is indeed stronger for short connections, and that long-distance connections serve more as a means of connecting local brain clusters, having evolved more recently as a result of cortical expansion. This hypothesis also suggests that the human brain, which has expanded even further than that of the macaque, should have an even higher proportion of such weak, long-range edges. While, due to its inherent biases, the use of probabilistic tractography to estimate edge weight from DWI data is problematic, it may nonetheless be inadvertently capturing real systematic bias towards stronger, shorter connections. Interestingly, Goulas et al. (2014) recently published a similar comparison of CoCoMac and DWI-based connectivity measures, using the same Regional Map parcellation as was utilized here. Comparing networks of similar density to the present study, the authors found significant correspondence between species on several topological measures. The geometric distribution of this correspondence (cf. Fig. 3 of that study) bears a qualitative resemblance to the average CI (i.e., across all modalities; Fig. 4D), as reported here.

The clustering results for R_{dwi} provide an interesting insight into its connectome (Figure S4). Specifically, inferior regions (medial and lateral temporal lobe and medial PFC) had poor interhemispheric symmetry, which likely contributed to the poor consensus clustering in these regions (Figure 7D). Both medial temporal lobe and medial PFC had a low mean CI, implying that these areas also have the worst binary correspondence (Figure 4D). This pattern is also evident when specifically considering only DMN regions (Figure 5). While there was general consensus for most superior ROIs in this network, there was much disagreement about inferior connections. In particular, R_{dwi} disagreed with all other modalities (including R_{cocomac}) in failing to detect a number of contralateral connections between inferior DMN regions. This disagreement may highlight a major limitation of DWI tractography, which is

the difficulty to navigate through, or terminate tracts in, regions of high uncertainty, such as those proximal to the corpus collosum. Despite the general improvement conferred by improved resolution in the NKI data acquisition, and the more computationally intensive methods used here, these results suggest that some fundamental issues in probabilistic tractography still remain. Further investigation of the behaviour of tractography algorithms in these particular regions may help resolve these issues.

4.8. fMRI network

Resting state fMRI captures functional synchrony in the absence of any overt task condition, which is proposed to represent "intrinsic" activations occurring in the absence of salient sensory input (reviewed in Rosazza and Minati, 2011; Kelly et al., 2012a). While this approach involves a high degree of individual variability, it has been consistently shown to activate regions of the DMN, which are typically suppressed in the presence of overt task demands. In the present study, we utilized this so-called iFC to compile an average connectome across subjects, which raises two clear issues. Firstly, while resting state fMRI typically produces both negative and positive correlations in individuals (as it did for our data set), for the averaged network, R_{fmri}, no negative correlations survived averaging across subsamples. This suggests that negative correlations were not as robust as positive, across all subjects, and that averaging may result in a loss of important information. While the group-wise approach was necessary for the present comparisons, further investigations could address this limitation by individually contrasting subject-wise connectomes from both fMRI and DWI tractography (e.g., as in Honey et al., 2009). Secondly, since no task-driven activation patterns are included in the present analysis, it is possible that some degree of FC – particularly of specialized modules – is not captured in the present comparison. However, a recent study by Smith et al. (2009), using a meta-analysis involving over 30,000 subjects, concluded that independent components derived from resting state fMRI could be matched to a wide array of task-based activation patterns, implying that iFC is sufficient to sample most of the brain's functional repertoire, even in the absence of overt task demands.

We found that R_{fmri} was highly symmetric across hemispheres, with very robust interhemispheric connectivity, a pattern that is well characterized in the literature (reviewed in Rosazza and Minati, 2011). As discussed above, however, a correlation in the activity of a pair of ROIs is not sufficient to support an inference about an effective connection between them, as it may arise from unconnected parallel pathways with common inputs, indirect connections with a third region (such as the thalamus), or temporal blurring inherent in the BOLD signal. There is some evidence that such mechanisms are indeed a factor. In humans, contralateral FC has been investigated in patients entirely lacking collosal fibres due to collosectomy, who showed only partial reductions in iFC (Uddin et al., 2008). Patients with agenesis of the corpus callosum (AgCC) have also been found to have reduced functional connectivity, measured by magnetoencepalography (MEG; Hinkley et al., 2012), but only in selective cortical regions; whereas a second, fMRI-based study of AgCC patients found similar iFC and clustering patterns to those of normal controls (Tyszka et al., 2011). While this evidence indicates that transcallosal connections have some relationship to interhemispheric FC, it also suggests that a high proportion of this FC is not necessarily mediated by direct anatomical connections. Contrasting R_{fmri} with other modalities (particularly R_{dwi} and R_{cocomac}) is thus an important step towards properly constraining a connectome for the human brain.

For fMRI, the issue of whether to perform global signal regression (GSR) is presently being debated in the neuroimaging community. The intention of GSR is to remove non-neuronal noise (i.e., physiological noise or motion artifacts), and thereby facilitate the detection of the neuronal signals of interest (Fox et al., 2009; Power et al., 2014). However, this preprocessing step may also introduce spurious negative correlations (Murphy et al., 2009), and could be removing important neuronal signals (Schölvinck et al., 2010; Fox et al., 2009). To assess the impact of GSR on the degree to which fMRI correlations correspond to other modalities, we performed GSR as an additional preprocessing step. Notably, our GSR approach differed from traditional approaches, as it involved regressing out the principal component most strongly correlated with the global signal, rather than that signal itself. This approach has been demonstrated not to produce spurious correlations (Carbonell et al., 2011). As shown in Figure 6, this GSR approach resulted in a substantial decrease in both weighted and binary agreement (ACC) with all three human

modalities, and for both the RM and Cambridge parcellations. The reasons for this decreased correspondence are not readily clear; while this result could indicate that GSR results in a worse representation of the underlying physical connectivity structure, it does not constitute definitive proof. An alternative possibility is that failing to remove GSR results in systematic non-neural correlations corresponding to DWI or structural covariance estimates which may themselves be partially artifactual or bias-driven.

Directly comparing fMRI connectivity estimates with and without GSR shows relatively moderate correlative agreement (mean $r^2 = 0.7$, leaving 30% unshared variance), considering they are derived from identical empirical signals (Figure 6A). Interestingly, when plotted across densities, it becomes clear that fMRI with GSR produces better binary correspondence at lower densities (see Figure 6B), which indicates that only at higher densities – with the involvement of less significant edges – do connectivity estimates produced by the GSR approach deviate substantially. This correspondence worsens for longer-distance connections, suggesting that longer edges are more altered by GSR preprocessing. These results appear to disagree with Fox et al. (2009), who found a tiny improvement of GSR over no GSR ($r = 0.36$ versus $r = 0.32$), comparing fMRI with DWI-based SC. However, since the SC (actually obtained from Honey et al. (2009)) was quite sparse in this study, the relative similarity in correspondence actually agrees with the binary correspondence found here for low density. Further insight is provided by the DMN renderings of the Cambridge parcellation networks (Figure 5B), which illustrate that the two preprocessing approaches produce networks with a similar general structure, but altered relative weights – with GSR producing stronger anteroposterior connectivity, and no GSR producing stronger local connectivity. These differences between weaker edges help illustrate the general trend of worse correspondence at higher densities for GSR-corrected fMRI, although the precise reason for such differences remains an open question.

4.9. Cortical parcellation

The cross-modal, cross-species comparison represented in this study is an ambitious project, given the plethora of issues associated with each modality, as well as more general model considerations. One of the larger difficulties is the choice of parcellation, both in terms of its interspecies correspondence and its particular granularity. We chose to use the Regional Map for two main reasons. Firstly, it was explicitly designed on the basis of widely-used cytoarchitectonic, gross anatomical, and functional criteria, in order to minimize cross-species discrepancies in ontology (Kötter and Wanke, 2005). Secondly, it has already been mapped across studies in the CoCoMac database, and therefore provides a convenient link between the monkey and human data. However, the granularity of this parcellation is relatively large (areas are on average $\sim 1600\ mm^2$) and its parcels are variable (the standard deviation of these areas is $\sim 1330\ mm^2$). Moreover, while the structural and functional landmark-based deformation of Van Essen (2004) is a good first approximation, cortical landmarks can be notoriously variable across humans (Amunts and Zilles, 2012), and this method makes an assumption of linear mapping between landmarks which may not reflect reality – especially for prefrontal regions (Petrides and Pandya, 2009).

Our choice of parcellation is therefore likely to have a substantial impact on the connectivity/covariance matrices we have obtained. To assess this further, we performed an additional comparison using the functionally-derived Cambridge template. While this parcellation has the disadvantage of not having a correspondence to the CoCoMac connectivity data, it also has a number of advantages: it is derived by voxel-wise clustering of resting-state fMRI across a large population, and thus reflects the natural (data-driven) separation of cortex based upon fMRI activations; and it has roughly twice as many regions, resulting in a much higher granularity. Interestingly, the use of this alternative parcellation resulted in a decrease of both correlative and binary correspondence (Figure 6), with $R_{\mathrm{dwi}}/R_{\mathrm{fmri}}$ (no GSR) having the best agreement ($r^2 = 0.27$). This reduced correspondence was somewhat unexpected, but may reflect the loss of smoothness inherent in decreasing parcel size. Alternately, increasing the number of parcels likely also increases the proportion of regional pairs with conflicting connectivity estimates. Notably, Honey et al. (2009) found a similar decrease in correspondence between RS-fMRI and DWI connectivity estimates (from $r = 0.66$ to $r = 0.36$; or $r = 0.82$ to $r = 0.48$ when considering a much sparser version which excluded absent SC), for 66 and 998 regions, respectively.

It is notable that under our current analysis we do not find strong resting-state functional covariance between PFCm and the more posterior CCp and CCr, despite this being a robust finding in RS-fMRI studies of the DMN (reviewed in van den Heuvel and Hulshoff Pol, 2010). While R_{dwi} shows strong connectivity between these regions, R_{fmri} connects more strongly to PFCpol (Figure 5). To investigate this discrepancy, we compared voxel-wise correlations across the brain for both the PFCm and CCp. As shown in Figure S3, we found significant voxel-wise anteroposterior correlations for both seed regions, but the strongest correlations occurred in the local vicinity of the seed region. This pattern was similar when ROIs were obtained from the functionally-derived Cambridge parcellation, and when using the standard (2500 ms) instead of short TR (645 ms) NKI images. When considering only the ROI-wise correlations that were used for our prior comparisons, we found that the strongest connectivity between CCr and prefrontal cortex was for region PFCpol, rather than PFCm. This was not the case with the GSR-corrected Cambridge parcellation, however – in this case, strong connectivity was found with both PFCm and PFCpol (Figure 5B). These considerations imply that correspondence across modalities depends to some extent on the granularity and specific border placement of the parcellation used, along with whether global regression was performed or not. The degree of scale dependence is thus an important question in the design of connectivity-based analyses, and we expect that emerging developments in multi-scale connectivity analysis can be adapted to the other modalities evaluated here – including the CoCoMac database – in order to address it.

4.10. Preprocessing

The present approach necessitated a series of specific preprocessing decisions, each of which is likely to have a nontrivial effect upon the resulting connectivity networks. For fMRI, the removal of motion and physiological artifacts, along with scanner artifacts, has produced many alternative techniques, each with its own set of caveats. While we did evaluate one of these options (GSR, above), the adoption of other pre-processing approaches (e.g., FIX; Salimi-Khorshidi et al., 2014) may have significant effects upon the correspondence with other modalities. For similar reasons as for fMRI, global signal regression is also an issue for structural covariance analyses, and remains largely unaddressed. DWI-based tractography methods are likewise subject to preprocessing conundrums. It is not certain, for instance, how or whether one might better correct tract counts for distance or anisotropy bias. The specific tractography algorithm used will also have a substantial impact on the resulting inferred networks (Bastiani et al., 2012). For these modalities, we chose not to explore the many possible pre-processing choices, but selected approaches which seemed most parsimonious according to our own standards. It is, however, important to note that our choices do not necessarily reflect consensus in the field. A useful extension of this study will be to assess the degree to which these preprocessing decisions affect cross-modal and cross-species correspondence.

4.11. Conclusions

The reasons for poor correspondence between connectomes derived from different imaging or experimental modalities are numerous, and include a number of methodological considerations – as highlighted above – as well as more conceptual ones. As an example of the latter, it is quite likely that differences between functional correlations and anatomical connections occur as a result of sequential activation by common input streams, which predict a temporal order that is not detected in the BOLD signal, or in the even more temporally blurred morphological covariance. This can be addressed by methods such as partial correlation, which attempts to quantify covariance after first factoring out the effect of competing ROIs (e.g., Smith et al., 2011), or by attempting to incorporate structural information in the statistical computation (an excellent Bayesian approach to this problem has recently been published by Hinne et al., 2013). BOLD acquisition methods which increase temporal resolution might also help temporally distinguish activation sequences (for instance, by facilitating temporal independent component analysis; Smith et al., 2012); indeed, this may partially account for the increased correspondence of R_{fmri} with R_{dwi} and $R_{\mathrm{thickness}}$ observed here for the NKI dataset. Ultimately, however, it is likely that more complex and nonlinear modelling approaches, which directly incorporate structural connectivity and functional activation or morphological change (e.g., Bojak et al., 2011), will be necessary to definitively construct a whole-brain human connectome from

neuroimaging data. The inclusion of EEG/MEG, which provide the temporal resolution that is not available from fMRI, will likely be indispensible in such approaches.

The present findings build upon similar cross-species comparisons reported in at least two other recent studies, one which compared CoCoMac to DWI tractography (Goulas et al., 2014), and another which compared macaque tractography to resting-state fMRI (Miranda-Dominguez et al., 2014). Here, we report a direct, whole-cortex comparison between macaque tract tracing evidence and connectivity inferred from multiple imaging modalities from a population of human subjects. This characterization serves a number of functions. Firstly, it cautions against equating the results of connectivity analyses drawn from different unimodal approaches. In other words, "connectivity" should not be used as an umbrella term; rather it should be understood as a phenomenon that cannot be directly observed, but only approximated in fundamentally different ways by a number of distinct methodologies. Secondly, it proposes a way by which to better understand and quantify the limitations peculiar to each modality, and by this means assess the efficacy of novel methodologies which attempt to address these limitations. Finally, it suggests that multimodal evidence for connectivity can be integrated in a manner which allows us to assign confidence levels to our inferences about particular connections, and thus get closer to an accurate representation of the effective human connectome.

4.12. Acknowledgements

We would like to acknowledge the help of Rembrandt Bakker in providing feedback, advice, and data from the CoCoMac database. We are also grateful to David Van Essen for very fruitful conversations about parcellation, interspecies registration, and other issues.

ATR was supported by a Canadian Institutes of Health Research (CIHR) Fellowship. SBE and ATR are supported by Deutsche Forschungsgemeinschaft (DFG, EI 816/4-1, LA 3071/3-1; EI 816/6-1.), the National Institute of Mental Health (R01-MH074457), the Helmholtz-Portfolio Project on "Supercomputing and Modeling for the Human Brain" and the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 604102 (Human Brain Project).

This research was enabled in part by support provided by Calcul Quebec (http://www.calculquebec.ca) and Compute Canada (http://www.computecanada.ca).

Amunts, K., Zilles, K., Aug 2012. Architecture and organizational principles of Broca's region. Trends Cogn. Sci. (Regul. Ed.) 16 (8), 418–426.

Baaré, W. F., Hulshoff Pol, H. E., Boomsma, D. I., Posthuma, D., de Geus, E. J., Schnack, H. G., van Haren, N. E., van Oel, C. J., Kahn, R. S., Sep 2001. Quantitative genetic modeling of variation in human brain morphology. Cereb. Cortex 11 (9), 816–824.

Bakker, R., Wachtler, T., Diesmann, M., 2012. CoCoMac 2.0 and the future of tract-tracing databases. Front Neuroinform 6, 30.

Bastiani, M., Shah, N. J., Goebel, R., Roebroeck, A., Sep 2012. Human cortical connectome reconstruction from diffusion weighted MRI: the effect of tractography algorithm. Neuroimage 62 (3), 1732–1749.

Behrens, T. E., Berg, H. J., Jbabdi, S., Rushworth, M. F., Woolrich, M. W., Jan 2007. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? Neuroimage 34 (1), 144–155.

Behrens, T. E., Sporns, O., Feb 2012. Human connectomics. Curr. Opin. Neurobiol. 22 (1), 144–153.

Bellec, P., June 2013. Mining the hierarchy of resting-state brain networks: Selection of representative clusters in a multiscale structure. In: Pattern Recognition in Neuroimaging (PRNI), 2013 International Workshop on. pp. 54–57.

Bellec, P., Carbonell, F. M., Perlbarg, V., Lepage, C., Lyttelton, O., Fonov, V., Janke, A., Tohka, J., Evans, A. C., 2011. A neuroimaging analysis kit for Matlab and Octave. In: Proceedings of the 17th International Conference on Functional Mapping of the Human Brain. pp. In Press+.

Bellec, P., Lavoie-Courchesne, S., Dickinson, P., Lerch, J. P., Zijdenbos, A. P., Evans, A. C., 2012. The pipeline system for Octave and Matlab (PSOM): a lightweight scripting framework and execution engine for scientific workflows. Front Neuroinform 6, 7.

Bellec, P., Rosa-Neto, P., Lyttelton, O. C., Benali, H., Evans, A. C., Jul 2010. Multi-level bootstrap analysis of stable clusters in resting-state fMRI. Neuroimage 51 (3), 1126–1139.

Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B (Methodological) 57 (1), 289–300. URL http://dx.doi.org/10.2307/2346101

Bezgin, G., Vakorin, V. A., van Opstal, A. J., McIntosh, A. R., Bakker, R., Aug 2012. Hundreds of brain maps in one atlas: registering coordinate-independent primate neuro-anatomical data to a standard brain. Neuroimage 62 (1), 67–76.

Bezgin, G., Wanke, E., Krumnack, A., Kötter, R., Oct 2008. Deducing logical relationships between spatially registered cortical parcellations under conditions of uncertainty. Neural Netw 21 (8), 1132–1145.

Bojak, I., Oostendorp, T. F., Reid, A. T., Kötter, R., Oct 2011. Towards a model-based integration of co-registered electroencephalography/functional magnetic resonance imaging data with realistic neural population meshes. Philos Trans A Math Phys Eng Sci 369 (1952), 3785–3801.

Büchel, C., Friston, K., 1997. Effective connectivity and neuroimaging. In: Frackowiak, R., Friston, K., Frith, C., Dolan, R., Mazziotta, J. (Eds.), Human Brain Function. Academic Press USA, pp. 127–140.

Carbonell, F., Bellec, P., Shmuel, A., 2011. Global and system-specific resting-state fMRI fluctuations are uncorrelated: principal component analysis reveals anti-correlated networks. Brain Connect 1 (6), 496–510.

Dauguet, J., Peled, S., Berezovskii, V., Delzescaux, T., Warfield, S. K., Born, R., Westin, C. F., Aug 2007. Comparison of fiber tracts derived from in-vivo DTI tractography with 3D histological neural tract tracer reconstruction on a macaque brain. Neuroimage 37 (2), 530–538.

Daunizeau, J., David, O., Stephan, K. E., Sep 2011. Dynamic causal modelling: a critical review of the biophysical and statistical foundations. Neuroimage 58 (2), 312–322.

Draganski, B., May, A., Sep 2008. Training-induced structural changes in the adult human brain. Behav. Brain Res. 192 (1), 137–142.

Eickhoff, S. B., Jbabdi, S., Caspers, S., Laird, A. R., Fox, P. T., Zilles, K., Behrens, T. E., May 2010. Anatomical and functional connectivity of cytoarchitectonic areas within the human parietal operculum. J. Neurosci. 30 (18), 6409–6421.

Fox, M. D., Zhang, D., Snyder, A. Z., Raichle, M. E., Jun 2009. The global signal and observed anticorrelated resting state brain networks. J. Neurophysiol. 101 (6), 3270–3283.

Friston, K. J., Frith, C. D., Liddle, P. F., Frackowiak, R. S., Jan 1993. Functional connectivity: the principal-component analysis of large (PET) data sets. J. Cereb. Blood Flow Metab. 13 (1), 5–14.

Geyer, S., Weiss, M., Reimann, K., Lohmann, G., Turner, R., 2011. Microstructural parcellation of the human cerebral cortex - From Brodmann's post-mortem map to in vivo mapping with high-field magnetic resonance imaging. Front Hum Neurosci 5, 19.

Giove, F., Gili, T., Iacovella, V., Macaluso, E., Maraviglia, B., Oct 2009. Images-based suppression of unwanted global signals in resting-state functional connectivity studies. Magn Reson Imaging 27 (8), 1058–1064.

Gong, G., He, Y., Chen, Z. J., Evans, A. C., Jan 2012. Convergence and divergence of thickness correlations with diffusion connections across the human cerebral cortex. Neuroimage 59 (2), 1239–1248.

Goulas, A., Bastiani, M., Bezgin, G., Uylings, H. B., Roebroeck, A., Stiers, P., Mar 2014. Comparative analysis of the macroscale structural connectivity in the macaque and human brain. PLoS Comput. Biol. 10 (3), e1003529.

Greicius, M. D., Supekar, K., Menon, V., Dougherty, R. F., Jan 2009. Resting-state functional connectivity reflects structural connectivity in the default mode network. Cereb. Cortex 19 (1), 72–78.

Hinkley, L. B., Marco, E. J., Findlay, A. M., Honma, S., Jeremy, R. J., Strominger, Z., Bukshpun, P., Wakahiro, M., Brown, W. S., Paul, L. K., Barkovich, A. J., Mukherjee, P., Nagarajan, S. S., Sherr, E. H., 2012. The role of corpus callosum development in functional connectivity and cognitive processing. PLoS ONE 7 (8), e39804.

Hinne, M., Ambrogioni, L., Janssen, R. J., Heskes, T., van Gerven, M. A., Oct 2013. Structurally-informed Bayesian functional connectivity analysis. Neuroimage.

Honey, C. J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J. P., Meuli, R., Hagmann, P., Feb 2009. Predicting human resting-state functional connectivity from structural connectivity. Proc. Natl. Acad. Sci. U.S.A. 106 (6), 2035–2040.

Kelly, C., Biswal, B. B., Craddock, R. C., Castellanos, F. X., Milham, M. P., Mar 2012a. Characterizing variation in the functional connectome: promise and pitfalls. Trends Cogn. Sci. (Regul. Ed.) 16 (3), 181–188.

Kelly, C., Toro, R., Di Martino, A., Cox, C. L., Bellec, P., Castellanos, F. X., Milham, M. P., Jul 2012b. A convergent functional architecture of the insula emerges across imaging modalities. Neuroimage 61 (4), 1129–1142.

Kelly, C., Uddin, L. Q., Shehzad, Z., Margulies, D. S., Castellanos, F. X., Milham, M. P., Petrides, M., Aug 2010. Broca's region: linking human brain functional connectivity data and non-human primate tracing anatomy studies. Eur. J. Neurosci. 32 (3), 383–398.

Kim, J. S., Singh, V., Lee, J. K., Lerch, J., Ad-Dab'bagh, Y., MacDonald, D., Lee, J. M., Kim, S. I., Evans, A. C., Aug 2005. Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification. Neuroimage 27 (1), 210–221.

Kötter, R., 2004. Online retrieval, processing, and visualization of primate connectivity data from the CoCoMac database. Neuroinformatics 2 (2), 127–144.

Kötter, R., Wanke, E., Apr 2005. Mapping brains without coordinates. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 360 (1456), 751–766.

Lerch, J. P., Worsley, K., Shaw, W. P., Greenstein, D. K., Lenroot, R. K., Giedd, J., Evans, A. C., Jul 2006. Mapping anatomical correlations across cerebral cortex (MACACC) using cortical thickness from MRI. Neuroimage 31 (3), 993–1003.

Liu, H., Stufflebeam, S. M., Sepulcre, J., Hedden, T., Buckner, R. L., Dec 2009. Evidence from intrinsic activity that asymmetry of the human brain is controlled by multiple factors. Proc. Natl. Acad. Sci. U.S.A. 106 (48), 20499–20503.

Lyttelton, O., Boucher, M., Robbins, S., Evans, A., Feb 2007. An unbiased iterative group registration template for cortical surface analysis. Neuroimage 34 (4), 1535–1544.

MacDonald, D., Kabani, N., Avis, D., Evans, A. C., Sep 2000. Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. Neuroimage 12 (3), 340–356.

Mantini, D., Corbetta, M., Romani, G. L., Orban, G. A., Vanduffel, W., Feb 2013. Evolutionarily novel functional networks in the human brain? J. Neurosci. 33 (8), 3259–3275.

Margulies, D. S., Vincent, J. L., Kelly, C., Lohmann, G., Uddin, L. Q., Biswal, B. B., Villringer, A., Castellanos, F. X., Milham, M. P., Petrides, M., Nov 2009. Precuneus shares intrinsic functional architecture in humans and monkeys. Proc. Natl. Acad. Sci. U.S.A. 106 (47), 20069–20074.

Markov, N. T., Ercsey-Ravasz, M., Van Essen, D. C., Knoblauch, K., Toroczkai, Z., Kennedy, H., Nov 2013. Cortical high-density counterstream architectures. Science 342 (6158), 1238406.

Markov, N. T., Ercsey-Ravasz, M. M., Ribeiro Gomes, A. R., Lamy, C., Magrou, L., Vezoli, J., Misery, P., Falchier, A., Quilodran, R., Gariel, M. A., Sallet, J., Gamanut, R., Huissoud, C., Clavagnier, S., Giroud, P., Sappey-Marinier, D., Barone, P., Dehay, C., Toroczkai, Z., Knoblauch, K., Van Essen, D. C., Kennedy, H., Sep 2012. A Weighted and Directed Interareal Connectivity Matrix for Macaque Cerebral Cortex. Cereb. Cortex 24 (1), 17–36.

Marrelec, G., Messé, A., Bellec, P., Jan. 2015. A Bayesian alternative to mutual information for the hierarchical clustering of dependent random variables. arXiv:1501.05194 [cs, q-bio, stat]ArXiv: 1501.05194.
URL `http://arxiv.org/abs/1501.05194`

Martinez, K., Madsen, S. K., Joshi, A. A., Joshi, S. H., Roman, F. J., Villalon-Reina, J., Burgaleta, M., Karama, S., Janssen, J., Marinetto, E., Desco, M., Thompson, P. M., Colom, R., Aug 2015. Reproducibility of brain-cognition relationships using three cortical surface-based protocols: An exhaustive analysis based on cortical thickness. Hum Brain Mapp 36 (8), 3227–3245.

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C., Collins, L., Thompson, P., MacDonald, D., Iacoboni, M., Schormann, T., Amunts, K., Palomero-Gallagher, N., Geyer, S., Parsons, L., Narr, K., Kabani, N., Le Goualher, G., Boomsma, D., Cannon, T., Kawashima, R., Mazoyer, B., Aug 2001. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). Philos. Trans. R. Soc. Lond., B, Biol. Sci. 356 (1412), 1293–1322.

McIntosh, A. R., Gonzalez-Lima, F., 1994. Structural equation modeling and its application to network analysis in functional brain imaging. Human Brain Mapping 2, 2–22.

Mechelli, A., Friston, K. J., Frackowiak, R. S., Price, C. J., Sep 2005. Structural covariance in the human cortex. J. Neurosci. 25 (36), 8303–8310.

Messé, A., Hutt, M. T., Konig, P., Hilgetag, C. C., 2015. A closer look at the apparent correlation of structural and functional connectivity in excitable neural networks. Sci Rep 5, 7870.

Miranda-Dominguez, O., Mills, B. D., Grayson, D., Woodall, A., Grant, K. A., Kroenke, C. D., Fair, D. A., Apr 2014. Bridging the gap between the human and macaque connectome: a quantitative comparison of global interspecies structure-function relationships and network topology. J. Neurosci. 34 (16), 5552–5563.

Mueller, S., Wang, D., Fox, M. D., Yeo, B. T., Sepulcre, J., Sabuncu, M. R., Shafee, R., Lu, J., Liu, H., Feb 2013. Individual variability in functional connectivity architecture of the human brain. Neuron 77 (3), 586–595.

Mukherjee, P., Chung, S. W., Berman, J. I., Hess, C. P., Henry, R. G., May 2008. Diffusion tensor MR imaging and fiber tractography: technical considerations. AJNR Am J Neuroradiol 29 (5), 843–852.

Murphy, K., Birn, R. M., Handwerker, D. A., Jones, T. B., Bandettini, P. A., Feb 2009. The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? Neuroimage 44 (3), 893–905.

Nooner, K. B., Colcombe, S. J., Tobe, R. H., Mennes, M., Benedict, M. M., Moreno, A. L., Panek, L. J., Brown, S., Zavitz, S. T., Li, Q., Sikka, S., Gutman, D., Bangaru, S., Schlachter, R. T., Kamiel, S. M., Anwar, A. R., Hinz, C. M., Kaplan, M. S., Rachlin, A. B., Adelsberg, S., Cheung, B., Khanuja, R., Yan, C., Craddock, C. C., Calhoun, V., Courtney, W., King, M., Wood, D., Cox, C. L., Kelly, A. M., Di Martino, A., Petkova, E., Reiss, P. T., Duan, N., Thomsen, D., Biswal, B., Coffey, B., Hoptman, M. J., Javitt, D. C., Pomara, N., Sidtis, J. J., Koplewicz, H. S., Castellanos, F. X., Leventhal, B. L., Milham, M. P., 2012. The NKI-Rockland Sample: A Model for Accelerating the Pace of Discovery Science in Psychiatry. Front Neurosci 6, 152.

Oguz, I., Farzinfar, M., Matsui, J., Budin, F., Liu, Z., Gerig, G., Johnson, H. J., Styner, M., 2014. DTIPrep: quality control of diffusion-weighted images. Front Neuroinform 8, 4.

Passingham, R., Feb 2009. How good is the macaque monkey model of the human brain? Curr. Opin. Neurobiol. 19 (1), 6–11.

Petrides, M., Pandya, D. N., Aug. 2009. Distinct Parietal and Temporal Pathways to the Homologues of Broca's Area in the Monkey. PLoS Biol 7 (8), e1000170.
URL http://dx.doi.org/10.1371/journal.pbio.1000170

Petrides, M., Tomaiuolo, F., Yeterian, E. H., Pandya, D. N., Jan 2012. The prefrontal cortex: comparative architectonic organization in the human and the macaque monkey brains. Cortex 48 (1), 46–57.

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., Petersen, S. E., Feb 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 59 (3), 2142–2154.

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., Petersen, S. E., Jan 2014. Methods to detect, characterize, and remove motion artifact in resting state fMRI. Neuroimage 84, 320–341.

Rilling, J. K., Glasser, M. F., Preuss, T. M., Ma, X., Zhao, T., Hu, X., Behrens, T. E., Apr 2008. The evolution of the arcuate fasciculus revealed with comparative DTI. Nat. Neurosci. 11 (4), 426–428.

Rilling, J. K., Seligman, R. A., May 2002. A quantitative morphometric comparative analysis of the primate temporal lobe. J. Hum. Evol. 42 (5), 505–533.

Roebroeck, A., Formisano, E., Goebel, R., Mar 2005. Mapping directed influence over the brain using Granger causality and fMRI. Neuroimage 25 (1), 230–242.

Rosazza, C., Minati, L., Oct 2011. Resting-state brain networks: literature review and clinical applications. Neurol. Sci. 32 (5), 773–785.

Salimi-Khorshidi, G., Douaud, G., Beckmann, C. F., Glasser, M. F., Griffanti, L., Smith, S. M., Apr 2014. Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. Neuroimage 90, 449–468.

Schölvinck, M. L., Maier, A., Ye, F. Q., Duyn, J. H., Leopold, D. A., Jun 2010. Neural basis of global resting-state fMRI activity. Proc. Natl. Acad. Sci. U.S.A. 107 (22), 10238–10243.

Sled, J. G., Zijdenbos, A. P., Evans, A. C., Feb 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Trans Med Imaging 17 (1), 87–97.

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., Filippini, N., Watkins, K. E., Toro, R., Laird, A. R., Beckmann, C. F., Aug 2009. Correspondence of the brain's functional architecture during activation and rest. Proc. Natl. Acad. Sci. U.S.A. 106 (31), 13040–13045.

Smith, S. M., Johansen-Berg, H., Jenkinson, M., Rueckert, D., Nichols, T. E., Miller, K. L., Robson, M. D., Jones, D. K., Klein, J. C., Bartsch, A. J., Behrens, T. E., 2007. Acquisition and voxelwise analysis of multi-subject diffusion data with tract-based spatial statistics. Nat Protoc 2 (3), 499–503.

Smith, S. M., Miller, K. L., Moeller, S., Xu, J., Auerbach, E. J., Woolrich, M. W., Beckmann, C. F., Jenkinson, M., Andersson, J., Glasser, M. F., Van Essen, D. C., Feinberg, D. A., Yacoub, E. S., Ugurbil, K., Feb 2012. Temporally-independent functional modes of spontaneous brain activity. Proc. Natl. Acad. Sci. U.S.A. 109 (8), 3131–3136.

Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., Ramsey, J. D., Woolrich, M. W., Jan 2011. Network modelling methods for FMRI. Neuroimage 54 (2), 875–891.

Stephan, K. E., Kamper, L., Bozkurt, A., Burns, G. A., Young, M. P., Kötter, R., Aug 2001. Advanced database methodology for the Collation of Connectivity data on the Macaque brain (CoCoMac). Philos. Trans. R. Soc. Lond., B, Biol. Sci. 356 (1412), 1159–1186.

Stephan, K. E., Zilles, K., Kötter, R., Jan 2000. Coordinate-independent mapping of structural and functional data by objective relational transformation (ORT). Philos. Trans. R. Soc. Lond., B, Biol. Sci. 355 (1393), 37–54.

Thirion, B., Varoquaux, G., Dohmatob, E., Poline, J. B., 2014. Which fMRI clustering gives good brain parcellations? Front Neurosci 8, 167.

Thompson, P. M., Cannon, T. D., Narr, K. L., van Erp, T., Poutanen, V. P., Huttunen, M., Lonnqvist, J., Standertskjold-Nordenstam, C. G., Kaprio, J., Khaledy, M., Dail, R., Zoumalan, C. I., Toga, A. W., Dec 2001. Genetic influences on brain structure. Nat. Neurosci. 4 (12), 1253–1258.

Tohka, J., Zijdenbos, A., Evans, A., Sep 2004. Fast and robust parameter estimation for statistical partial volume models in brain MRI. Neuroimage 23 (1), 84–97.

Tyszka, J. M., Kennedy, D. P., Adolphs, R., Paul, L. K., Oct 2011. Intact bilateral resting-state networks in the absence of the corpus callosum. J. Neurosci. 31 (42), 15154–15162.

Uddin, L. Q., Mooshagian, E., Zaidel, E., Scheres, A., Margulies, D. S., Kelly, A. M., Shehzad, Z., Adelstein, J. S., Castellanos, F. X., Biswal, B. B., Milham, M. P., May 2008. Residual functional connectivity in the split-brain revealed with resting-state functional MRI. Neuroreport 19 (7), 703–709.

Urchs, S., Dansereau, C., Benhajali, Y., Bellec, P., 01 2015. Group multiscale functional template generated with BASC on the Cambridge sample.
URL http://dx.doi.org/10.6084/m9.figshare.1285615

van den Heuvel, M. P., Hulshoff Pol, H. E., Aug 2010. Exploring the brain network: a review on resting-state fMRI functional connectivity. Eur Neuropsychopharmacol 20 (8), 519–534.

Van Dijk, K. R., Hedden, T., Venkataraman, A., Evans, K. C., Lazar, S. W., Buckner, R. L., Jan 2010. Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. J. Neurophysiol. 103 (1), 297–321.

Van Essen, D. C., 2004. Surface-based approaches to spatial localization and registration in primate cerebral cortex. Neuroimage 23 Suppl 1, 97–107.

Van Essen, D. C., Nov 2005. A Population-Average, Landmark- and Surface-based (PALS) atlas of human cerebral cortex. Neuroimage 28 (3), 635–662.

Van Essen, D. C., Dierker, D. L., Oct 2007. Surface-based and probabilistic atlases of primate cerebral cortex. Neuron 56 (2), 209–225.

Ward, J. H., 1963. Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association 58 (301), 236–244.
URL http://www.jstor.org/stable/2282967

Wright, I. C., Sham, P., Murray, R. M., Weinberger, D. R., Bullmore, E. T., Sep 2002. Genetic contributions to regional variability in human brain structure: methods and preliminary results. Neuroimage 17 (1), 256–271.

Zijdenbos, A. P., Forghani, R., Evans, A. C., Oct 2002. Automatic "pipeline" analysis of 3-D MRI data for clinical trials: application to multiple sclerosis. IEEE Trans Med Imaging 21 (10), 1280–1291.

Zilles, K., Armstrong, E., Schleicher, A., Kretschmann, H. J., 1988. The human pattern of gyrification in the cerebral cortex. Anat. Embryol. 179 (2), 173–179.