

## Supplementary Results and Discussion For A Multi-modal Parcellation of Human Cerebral Cortex

Matthew F. Glasser<sup>1</sup>, Timothy S. Coalson<sup>1\*</sup>, Emma C. Robinson<sup>2,3\*</sup>, Carl D. Hacker<sup>4\*</sup>, John Harwell<sup>1</sup>, Essa Yacoub<sup>5</sup>, Kamil Ugurbil<sup>5</sup>, Jesper Andersson<sup>2</sup>, Christian F. Beckmann<sup>6</sup>, Mark Jenkinson<sup>2</sup>, Stephen M. Smith<sup>2</sup>, David C. Van Essen<sup>1</sup>

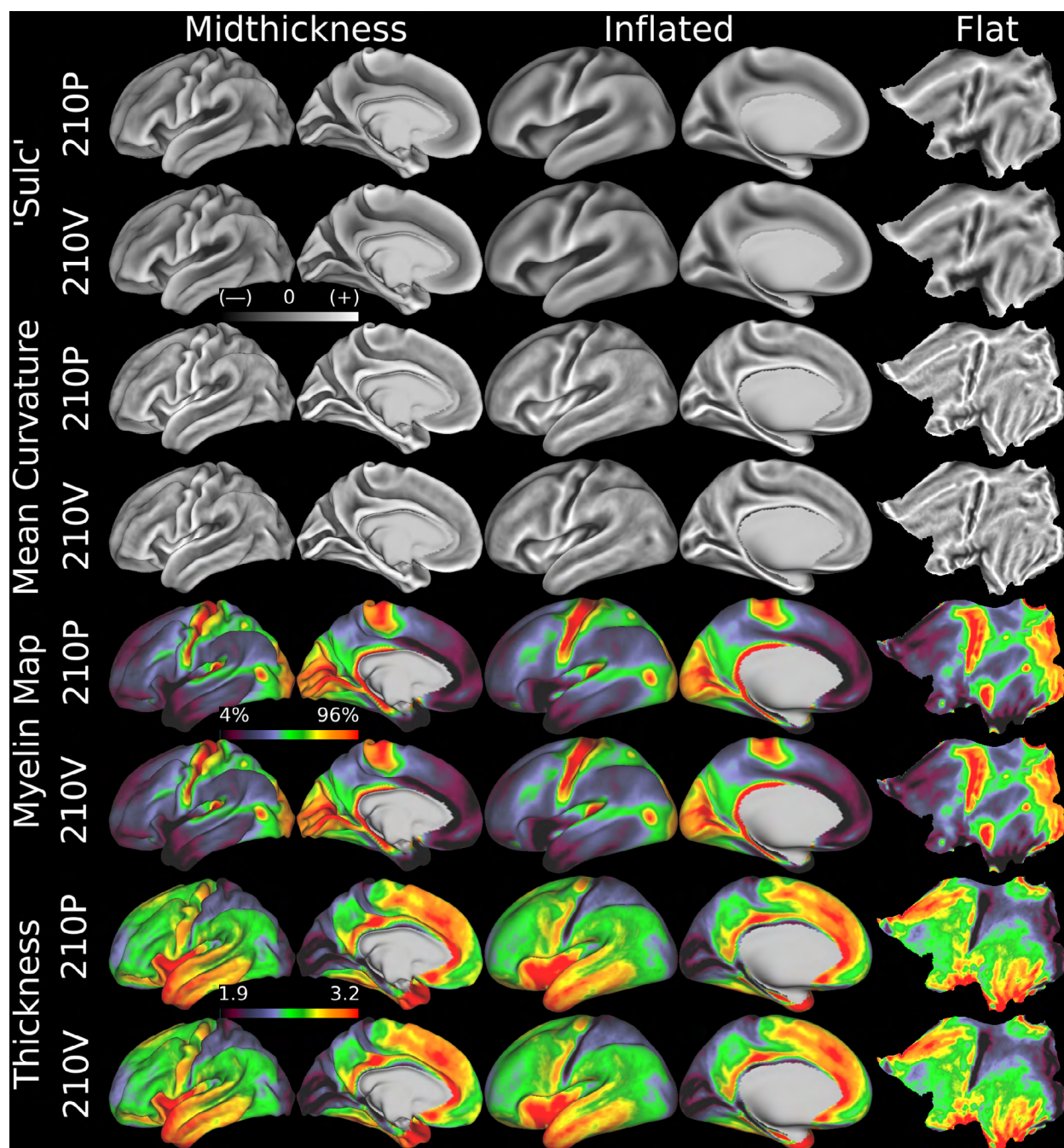
<sup>1</sup>*Department of Neuroscience, Washington University Medical School, Saint Louis, Missouri 63110, USA.* <sup>2</sup>*FMRIB Centre, Nuffield Department of Clinical Neurosciences, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DU, UK.* <sup>3</sup>*Department of Computing, Imperial College, London SW7 2AZ, UK.* <sup>4</sup>*Department of Biomedical Engineering, Washington University, Saint Louis, Missouri 63110, USA.* <sup>5</sup>*Center for Magnetic Resonance Research (CMRR), University of Minnesota, Minneapolis, Minnesota 55455, USA.* <sup>6</sup>*Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen 6525 EN, The Netherlands.* <sup>7</sup>*Department of Cognitive Neuroscience, Radboud University Medical Centre Nijmegen, Postbus 9101, Nijmegen 6500 HB, The Netherlands.*

*\*These authors contributed equally to this work.*

### 1. Supplementary Results

#### 1.1 Reproducibility of group average multi-modal data

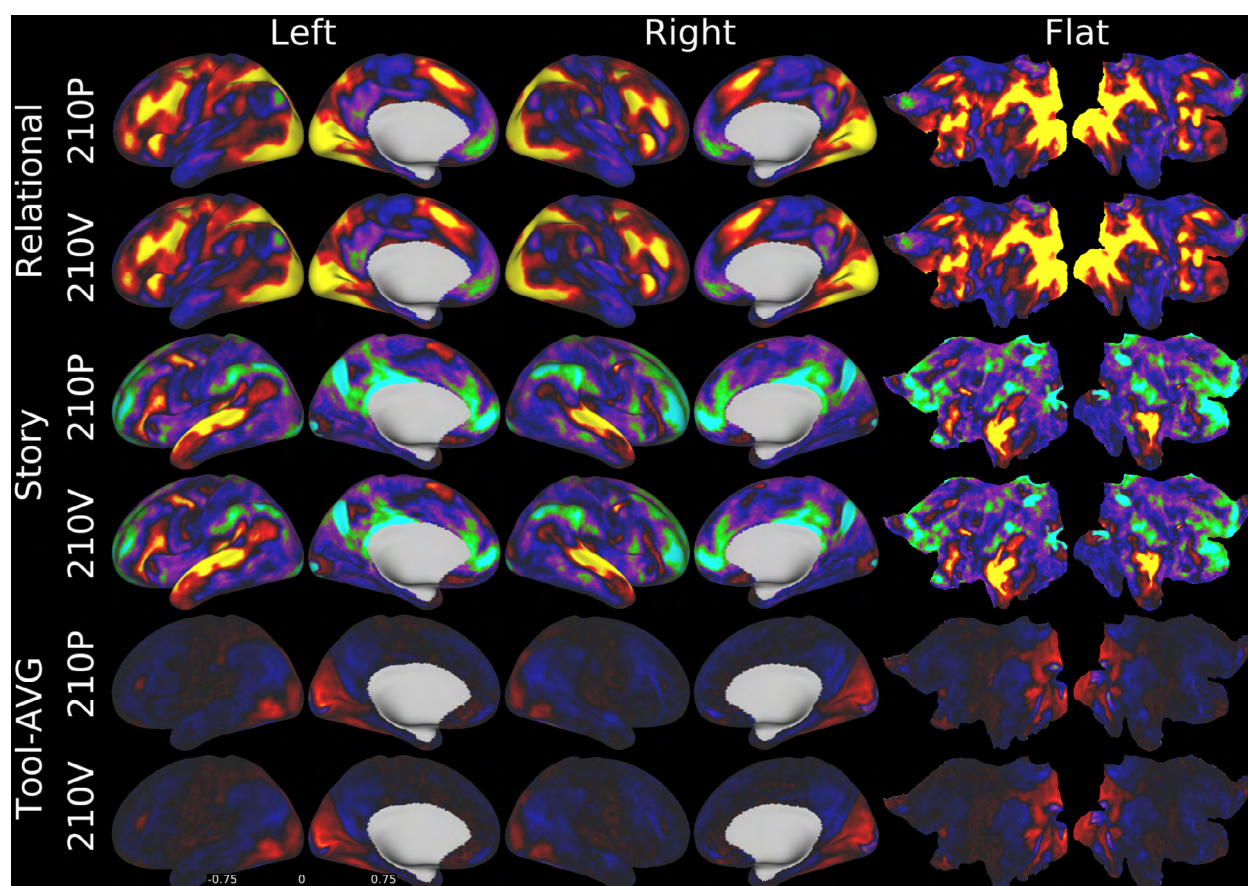
Expanding on the striking similarities between 210P and 210V datasets highlighted in Main Text Figure 1, Figure 1 shows group average midthickness surfaces, FreeSurfer sulc and curvature folding measures, myelin maps, and cortical thickness maps corrected for surface curvature. All of these measures are highly reproducible across the 210P and 210V group averages (see legend to Figure 1). Because MSMAll areal feature-based registration was used to register the data (using myelin maps, resting state network maps, and visuotopic maps, see Methods section on image preprocessing), folding features that are well correlated with cortical areas remain sharp, whereas folding features that are poorly correlated with cortical areas are blurred away. This results in sharp folding patterns for the central sulcus, calcarine sulcus, and the insula, where areal boundaries are closely related to folds, whereas the folding patterns are blurred out in more cognitive regions where areas and folds are not well as correlated (see Figure 7 below for a side by side comparison of group and individual folding maps). This blurring effect on folding patterns in many cortical regions is also visible in the group average midthickness surfaces of Figure 1. Fine details in the average myelin maps and thickness maps are preserved because these features are correlated more strongly with cortical areas.



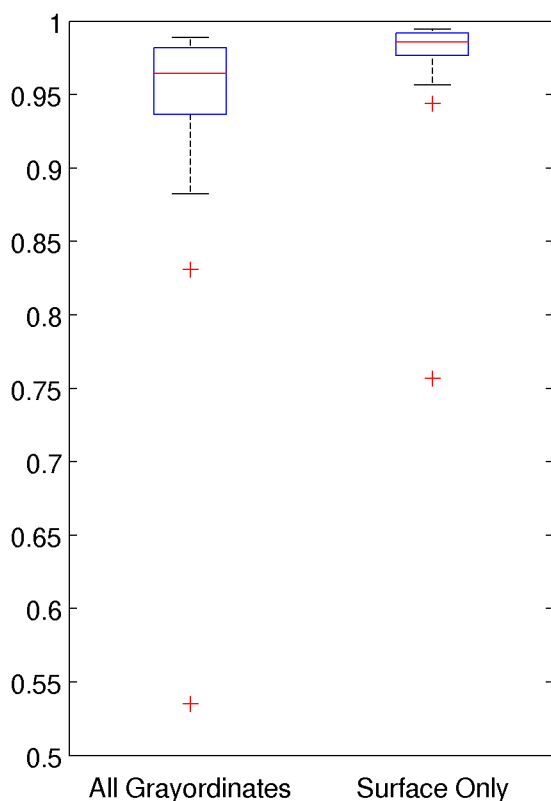
**Figure 1 | HCP group average architectural and folding map reproducibility.** Here we show group average left midthickness surfaces in Columns 1 and 2, left inflated surfaces in Columns 3 and 4 and left flat maps in Column 5. The 210P group averages are the odd rows and the 210V group averages are the even rows. The top two rows are FreeSurfer 'sulc' ( $r=0.996$ ), the next two are FreeSurfer 'curv' ( $r=0.979$ ), the next two are myelin ( $r=0.998$ ), and the final two are thickness corrected for folding effects ( $r=0.994$ ); all show extremely high cross-group reproducibility (both hemispheres were used in computing spatial map correlation coefficients). Folding patterns (sulc and curv) become blurry in many higher association regions of cortex, though they remain sharp in primary regions such as the central sulcus, calcarine sulcus, and insula, indicating that these regions have cortical areas that are well correlated with folding patterns. Areal feature-based surface registration also preserves fine spatial details in myelin maps and thickness maps that are consistent across groups. An example is the reproducible variation in the myelin maps of the primary somatosensory cortex, which appears to be related to somatotopic subdivisions (see Supplementary

Neuroanatomical Results #6, where these variations are compared with functional and connectivity-based somatotopic subdivisions). Sulcal folding is dark and gyral folding is bright, with grey in between. High myelin and thick cortex are red, low myelin and thin cortex are dark. T1w/T2w myelin content maps are a relative measure depicted on a percentile scale that indicates which cortical areas have more or less myelin. For brevity, we refer to them as ‘myelin maps’ in this report (as in previous reports). We consider them akin to an *in vivo*, MRI-based “myelin stain,” analogous to histological myelin stains long used by neuroanatomists rather than a truly quantitative measure of intracortical myelin (Glasser et al., 2014; Glasser and Van Essen, 2011). Data at <http://balsa.wustl.edu/jpM6>.

Figure 2 shows the spatial detail and quantitative reproducibility achievable without using spatial smoothing of the effect size (beta) maps for three example task fMRI contrasts (highest reproducibility, median reproducibility, least reproducible; Z statistic maps are similarly reproducible). Figure 3 shows the reproducibility for all 86 task contrasts, both for all grayordinates and for just surface grayordinates.

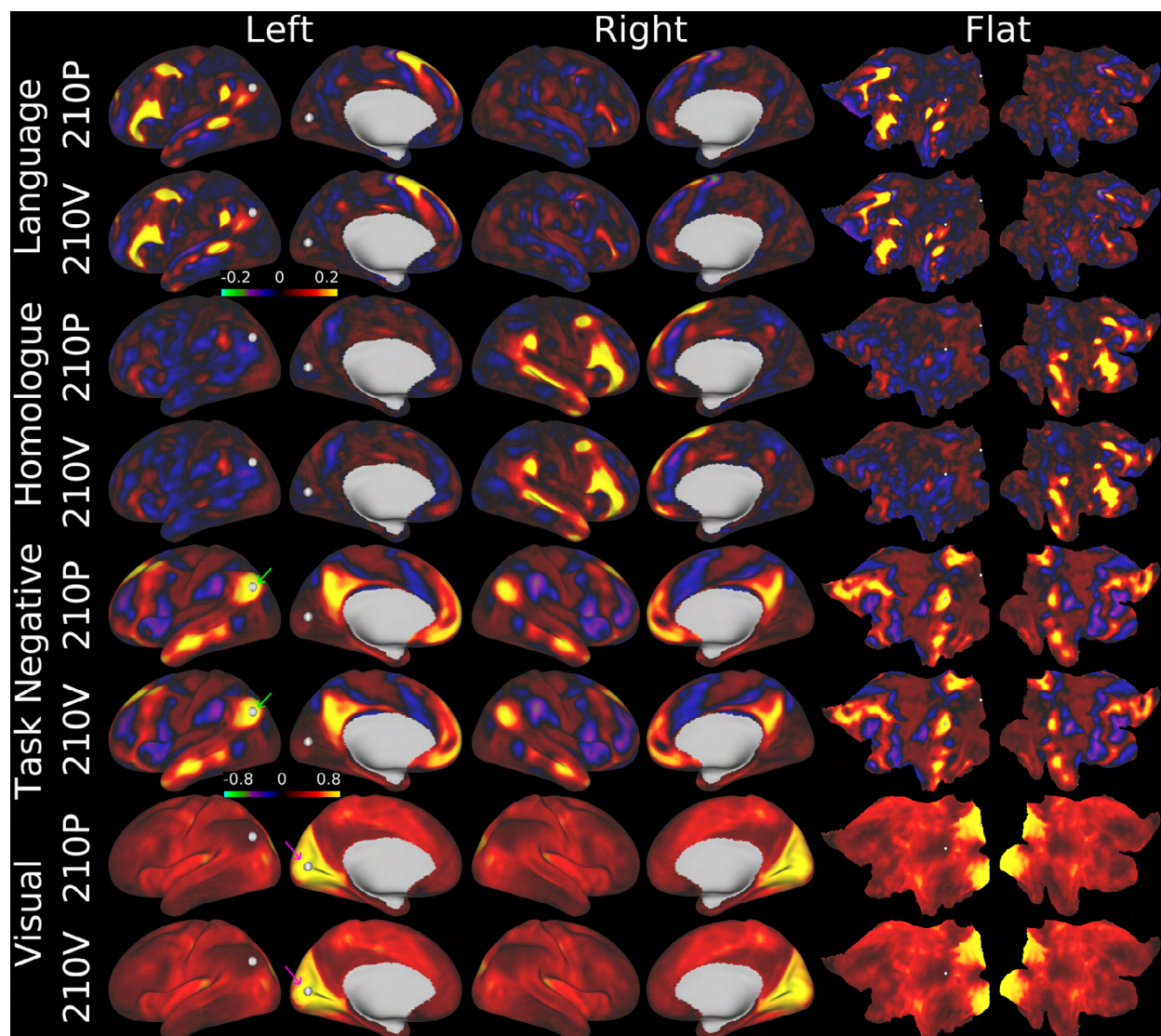


**Figure 2 | HCP group average task fMRI reproducibility.** Images are scaled from beta = +/-0.75. The odd rows are from the 210P dataset whereas the even rows are from the 210V dataset. These data are shown on left inflated views (Columns 1 and 2), right inflated views (Columns 3 and 4), and left and right flat maps (Columns 5 and 6, respectively). The top two rows are the task fMRI contrast with the highest reproducibility (Relational vs Baseline:  $r=0.995$ ), the middle two rows are the task fMRI contrast with the median reproducibility (Language: Story vs Baseline:  $r=0.984$ ), and the last two rows are the task fMRI contrast with the lowest reproducibility (Working Memory/Categories: Tool-Avg:  $r=0.944$ ) that is not a known outlier (see Figure 3 legend). See (Robinson et al., 2014) for a comparison of the sharpness of task fMRI maps generated with areal feature-based registration versus folding-based registration. Data at <http://balsa.wustl.edu/Q94z>.

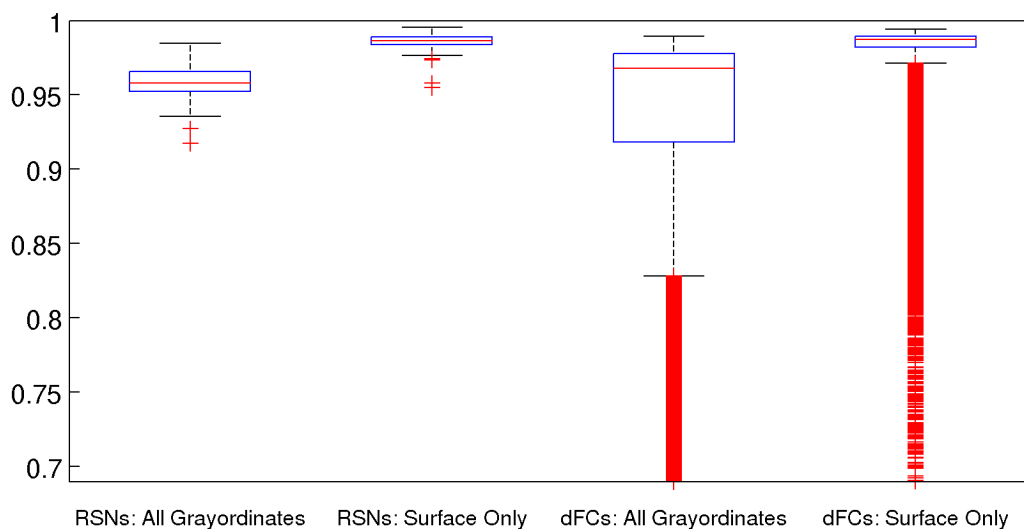


**Figure 3 | HCP group average task fMRI reproducibility for all contrasts.** Here we show box and whisker plots demonstrating the reproducibility of all 86 task contrasts (47 unique, 39 sign reversed) determined for all grayordinates (left) or surface grayordinates only (right, as shown in Figure 2 and whose r-values are mentioned in the figure legend). The bottom outlier (red cross) is the Gambling Reward-Punishment task contrast, which did not work as intended by the HCP and has very low information content in individual subjects. See Figure 2 legend for corresponding image data.

The top four rows of Figure 4 show the spatial patterns of two different resting state network maps (RSNs) from a  $d=137$  ICA decomposition after computation of the individual subject maps using weighted regression and then averaging across the 210P group and the 210V group. (Weighted regression produces more accurate individual subject maps than the more standard dual regression approach, see Supplementary Methods #2.3.) The bottom four rows of Figure 4 show the reproducibility of dense functional connectivity maps for two different seeds. Figure 5 shows the reproducibility of all the non-artifactual RSNs and for all dense functional connectivity maps. Using the  $d=130$  RSNs computed from the 210V subjects yielded similar reproducibility numbers (data not shown). The method for choosing the ICA dimensionalities is explained in Supplementary Methods #3.3.



**Figure 4 | HCP group average resting-state fMRI reproducibility.** Here we show the reproducibility of Resting State Networks (RSNs, ICA  $d=137$ ) and dense Functional Connectivity maps. The odd rows are the 210P dataset and the even rows are the 210V dataset. The RSNs were produced using ICA on the 210P dataset and then weighted regression of the group spatial maps onto the individual subject dense timeseries to produce individual subject spatial maps in all 449 subjects. The individual subject maps were then separately averaged across the 210P and 210V groups. The top two rows show the left lateralized language network and the next two rows show its right hemisphere homologue, both scaled between  $\beta = \pm 0.2$ . The next two rows show dense functional connectivity (FC) of the bilateral task negative (default mode) network seeded from a posterior inferior parietal grayordinate (green arrow, white sphere on left lateral surface), which has strong anti-correlation with the task positive network. The last two rows show the early visual network seeded from a grayordinate in the center of the calcarine sulcus (purple arrow, white sphere on left medial surface), which shows positive correlation with the entire brain. The dense FC maps are scaled from  $r = -0.8$  to  $0.8$ . Data at <http://balsa.wustl.edu/RkMg>.



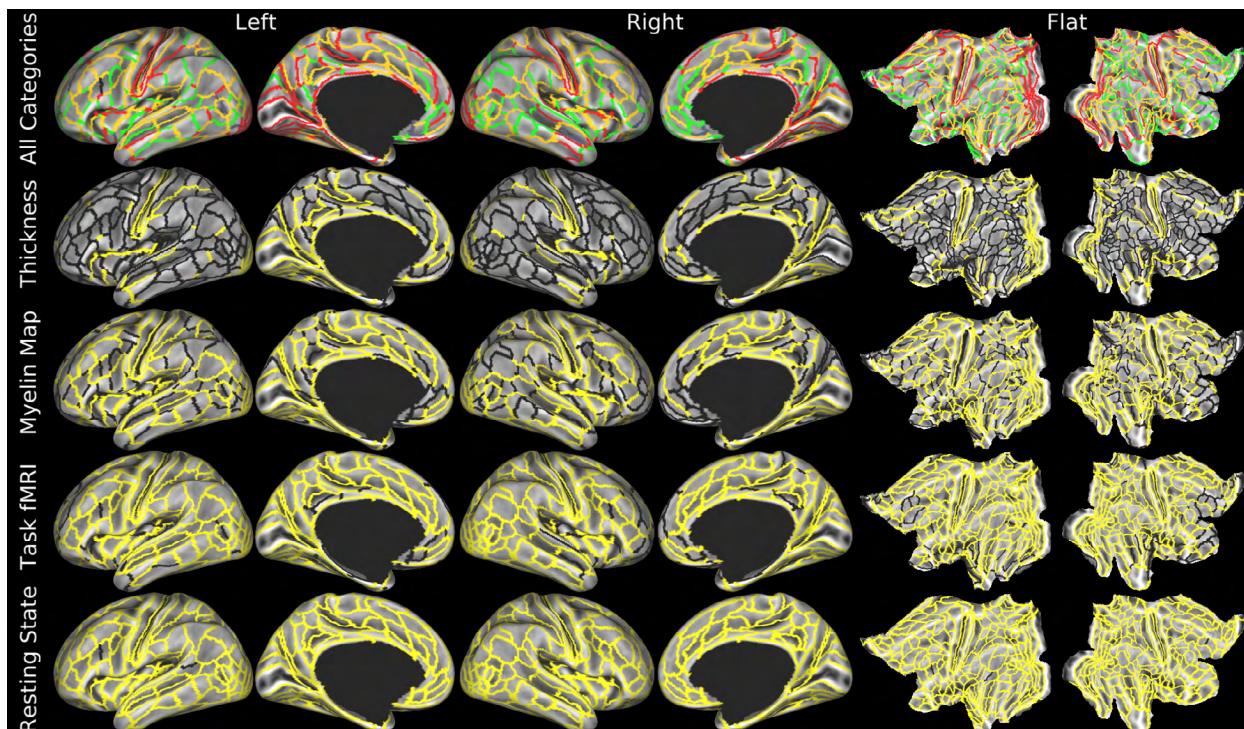
**Figure 5 | HCP group average resting state fMRI reproducibility all data.** Here we show the reproducibility of the 210P and 210V subject groups of resting state network maps (RSNs) ( $d=137$ ) and dense Functional Connectivity maps (dFCs) for all grayordinates and for only surface grayordinates. Group level artifactual components (3) are not included ( $\#RSNs=134$ ) for either measure and only primarily surface components ( $\#RSNs=77$ ) are included for the surface only measure. See Figure 4 legend for how the RSNs were produced. For RSNs, the all grayordinates median reproducibility is  $r=0.978$  and for surface only it is  $r=0.989$ . For dense FC maps, the all grayordinates median reproducibility is  $r=0.973$  and for surface only it is  $r=0.989$ . Dense FC maps contain a considerable number of outliers (e.g. seeds in low CNR grayordinates in cortical and subcortical susceptibility regions). For dense FC, some outliers (red pluses more than 2.7 standard deviations below the median) extend below  $r=0.7$ , however the 134 RSN maps contain few outliers (as they have values largely close to zero in these low CNR susceptibility regions). Some of the dense map outliers may reflect the very large number of samples (over 4 billion) in a dense connectome. See Figure 4 for corresponding image data.

### 1.2 Cross-validation of the multi-modal parcellation

For the statistical cross-validation, the 210V dataset was parcellated using the 210P-derived multi-modal parcellation (i.e., architectural measures and fMRI timeseries for each 210V subject were averaged within the 210P-defined parcels prior to functional or connectivity-based analyses). We then computed paired t-tests (for significance) and effect sizes (Cohen's  $d$ , for robustness) across all subjects for each pair of neighboring areas and for all 266 distinct feature maps within the four independent feature categories: cortical thickness, myelin content, task fMRI contrast map betas, and full correlation functional connectivity with the other 178 areas (i.e. excluding the diagonal) (see Supplementary Methods, #7.2). To determine whether each of these parcellated feature maps was indeed robustly and statistically significantly different across the area pair's border in individual subjects, the results were thresholded at a Cohen's  $d > 1$  effect size threshold and a statistical significance threshold of  $p < 9 \times 10^{-8}$  (Bonferroni-corrected for both hemispheres and two tails, see Methods section on statistical cross validation).

Figure 6 shows results displayed at each of the 1,050 borders shared by a pair of areas in the left or the right hemisphere. Colors in the top row indicate that 205 areal

borders are robustly and statistically significantly different across all four feature categories (red, 19.5% of border pairs), 535 across three categories (yellow, 51.0%), 268 across two categories (green, 25.5%), and 40 across one category (blue, 3.8%). Only two areal borders (black, 0.2%) were below the robust effect size threshold for all feature categories (but nonetheless passed the statistical significance threshold for multiple categories, see Figure 6 legend). Rows 2 - 5 show which areal borders passed the thresholds for each feature category. These results demonstrate that the areas of the multi-modal parcellation differ from their neighbors across multiple modalities in an independent group of subjects.



**Figure 6 | Cross-validation of the multi-modal parcellation.** Here we show the results of paired t-tests between each pair of neighboring areas' mean values across subjects in the 210V dataset but using the borders from the 210P parcellation. These tests were thresholded at Cohen's  $d > 1$  ( $\text{abs}(\text{mean paired difference}) / \text{standard deviation of paired differences}$ ) and  $p < 9 \times 10^{-8}$  for both hemispheres (Bonferroni corrected across two hemispheres, two-tailed significance threshold, see Methods section on statistical cross validation) to identify robust and statistically significant differences across each border pair. Row 1 shows how many of the four categories of features (thickness, myelin, tfMRI, rfMRI) passed these thresholds (Red=4, Orange=3, Green=2, Blue=1, Black=0). Rows 2 - 5 show which borders passed the thresholds (yellow=passed, black=not passed) respectively for thickness, myelin, any task fMRI contrast, and any difference in connectivity with an area in the full correlation parcellated connectome (excluding the diagonal). Two borders do not pass the  $d > 1$  effect size threshold for any category but still pass the significance threshold for multiple modalities (Black in Row 1): IFJa/IFSp in the left hemisphere (myelin:  $d = 0.91$ ,  $p = 3 \times 10^{-29}$ , tfMRI most different contrast LANGUAGE Story-Math:  $d = 0.62$ ,  $p = 2 \times 10^{-16}$ , rfMRI most different connectivity with area 47m:  $d = -0.85$ ,  $p = 9 \times 10^{-27}$ ) and 9a/9p in the right hemisphere (myelin:  $d = -0.84$ ,  $p = 4 \times 10^{-26}$ , rfMRI most different connectivity with area 8Ad:  $d = -0.56$ ,  $p = 3 \times 10^{-14}$ ). Data at <http://balsa.wustl.edu/WON5>.

### 1.3 Typical and atypical areal topologies of 55b and its neighbors

Although area 55b is a notably well defined cortical area (see Main Text Figure 2), it also exhibits a striking degree of individual variability in location and shape, even after MSMAll areal-feature-based surface registration. A substantial minority of subjects have topologically incompatible organizations of 55b and surrounding areas and areal features. These atypical topologies include spatial shifts, in which 55b is shifted relative to nearby areas and features, and splits of 55b, in which 55b is bisected by another area. Figure 7 illustrates spatial shifts in area 55b relative to another major nearby feature, the upper limb somatosensory-motor subregion. Fig. 7a-e show the group average areal features and pattern, in which 55b is strongly connected with the language network (RSN in column 1, task in column 3), is located between more heavily myelinated areas FEF, PEF, and 4 (column 2), and is elongated along an axis that points towards the postero-infero-lateral tip of the upper limb subregion of the somato-motor strip (RSN in column 4). Fig. 7f-j show a typical individual subject whose areal features closely match those of the group average. Fig. 7k-t show two subjects with an atypical organization in and around area 55b. Instead of 55b's long axis pointing ventrally, in both subjects it is tilted up towards the supero-anterior portion of the upper limb region, actually contacting it at the white sphere (columns 1 and 4). Although MSMAll attempts to align all of these areal features, the much larger upper limb region "wins," and hence area 55b remains misaligned with the group average. Because the patterns are topologically incompatible, MSMAll would generate large distortions in this region if the regularization penalty were relaxed. Column 5 shows the corresponding folding maps for the group average and individual subjects, which show both substantial variability and relatively poor correlation between areal boundaries and folding patterns in the region anterior to the central sulcus.

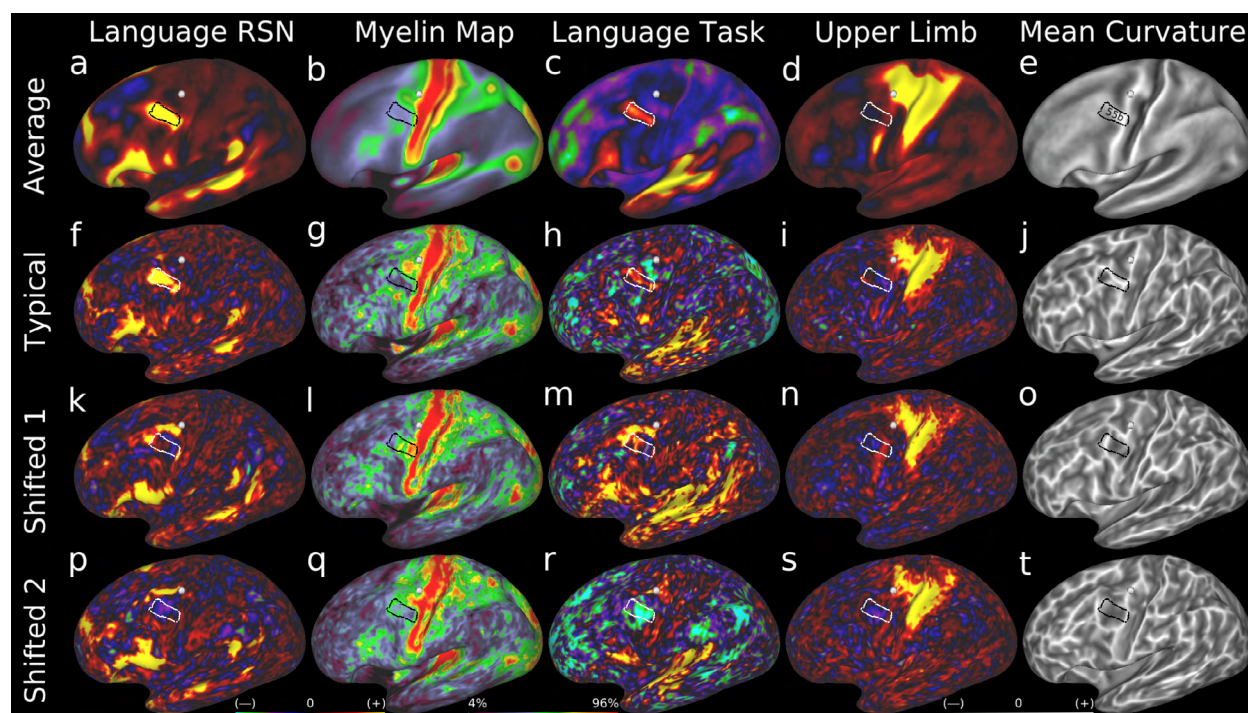


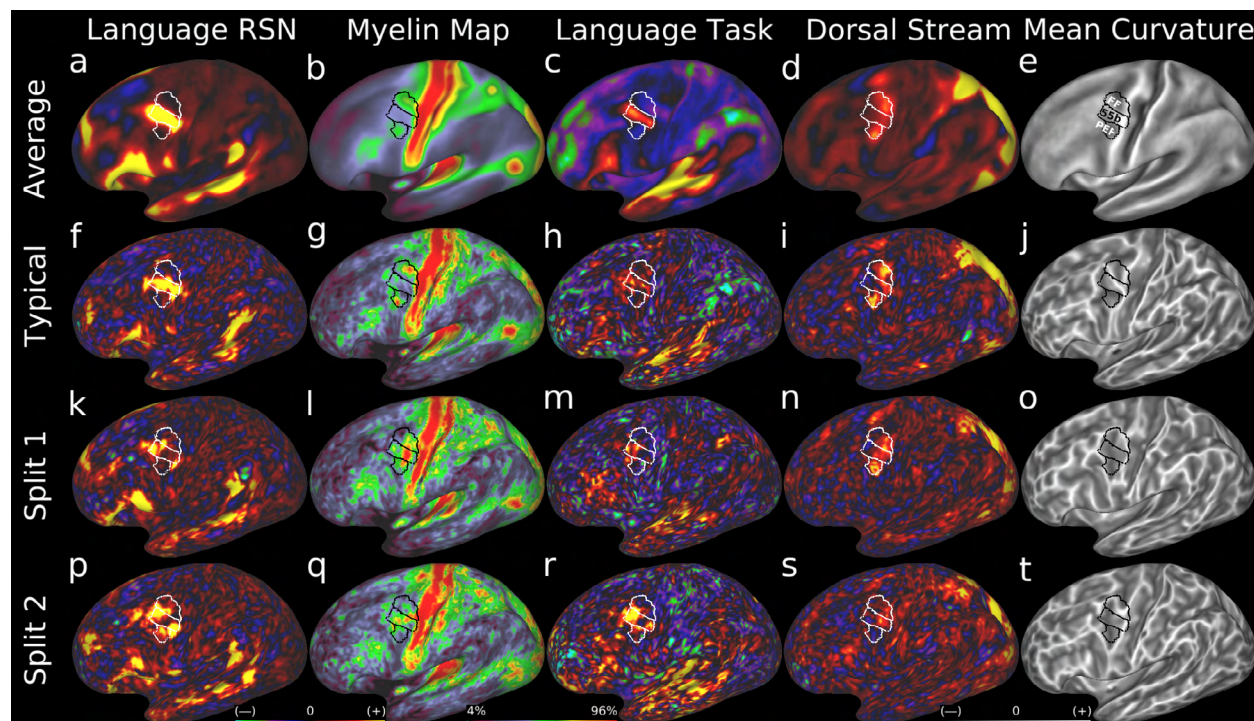


Figure 7 | **Atypical shifts of area 55b.** Here we show examples of one type of topological incompatibility between the typical layout of areal features near area 55b and a minority of subjects with an alternative ‘shifted’ layout. The first row shows the group average areal features: **a**, a  $d=40$  RSN map corresponding to parts of the language network, **b**, a myelin map, **c**, a LANGUAGE Story-Baseline task contrast beta map, **d**, a  $d=40$  RSN map corresponding to the upper limb somatosensory-motor topographic subregion, and, **e**, a mean curvature map illustrating the folding pattern. The group delineation of area 55b is a white or black line. **f-j** show these maps in a typical individual subject. **k-o** show an atypical subject whose 55b areal features are shifted relative to the typical pattern. **p-t** show a second atypical subject with a very similar pattern. In both cases, the posterior portion of area 55b is located near the white sphere instead of lying within the group delineated area 55b. Data at <http://balsa.wustl.edu/Q2LV>.

Besides the atypical shifted pattern just illustrated, area 55b and adjoining areas FEF and PEF show an additional atypical pattern. Typically, FEF and PEF are two moderately heavily myelinated areas that share very similar functional activation and connectivity patterns, but are split by lightly myelinated area 55b, which has very different functional activation and connectivity patterns. In a minority of subjects, however, it is 55b that is split, and FEF and PEF merge so that they are spatially adjoining. Figure 8 shows examples of this alternative topology. Fig. 8a-e show the group average areal features and pattern, in which 55b is strongly connected with the language network and is located between more heavily myelinated areas FEF, PEF, and 4. Also FEF and PEF are connected with the dorsal visual stream (RSN in column 4). Fig. 8f-j show a typical individual subject whose areal features closely match those of the group average. Fig. 8k-t show two subjects having the split 55b, joined FEF and PEF pattern. Importantly, the gap within the expected location of 55b in columns 1 and 3 (Fig. 8k, 8p, 8m, & 8r), indicating a lack of association (connectivity and function) with the language network, aligns with the region of increased connectivity with the dorsal stream visual cortex (column 4, Fig. 8n & 8s) and with elevated myelin content (column 2, Fig. 8l & 8q). The multi-modal nature of these individual differences argues strongly against attributing these findings simply to “noise in the data”; instead, the evidence indicates genuine individual differences in the topology of cortical functional and structural organization. As with the shifted patterns, the folding patterns in and near 55b, FEF, and PEF are both variable across subjects and variable in their relationship to areal features.

The neuroanatomists’ survey of the left hemisphere of the 210P group of subjects using these multi-modal areal features revealed that ~89% ( $n=186$ ) have the typical topology, ~4% ( $n=9$ ) have the shifted topology, ~6% ( $n=12$ ) have the split topology, and ~1% ( $n=3$ ) have some other topology. While the canonical parcellation is best defined on the basis of the typical topology, it is highly desirable to map the alternative areal topologies as well as possible so that functional or connectivity-based measures are comparing the same areas across subjects. This is a major reason for using an areal classifier to delineate individual subject parcellations instead of simply imposing an atlas parcellation on the data. No topology preserving registration algorithm can bring these atypical subjects into precise alignment with an atlas parcellation. We found that area 55b is a notable hotspot of variability in areal topology (that is correlated with a hotspot in areal distortion from the MSMAll registration). Such variability in areal topologies will likely occur in varying degrees throughout the cortex, however. Presumably, as one examines finer levels of the neural hierarchy (e.g. within-area modularity), analogous kinds of spatially incompatible topologies will be increasingly prevalent. To take an extreme, it is

presumably not possible to register brains at the level of individual neurons or even cortical columns in a topology preserving way.

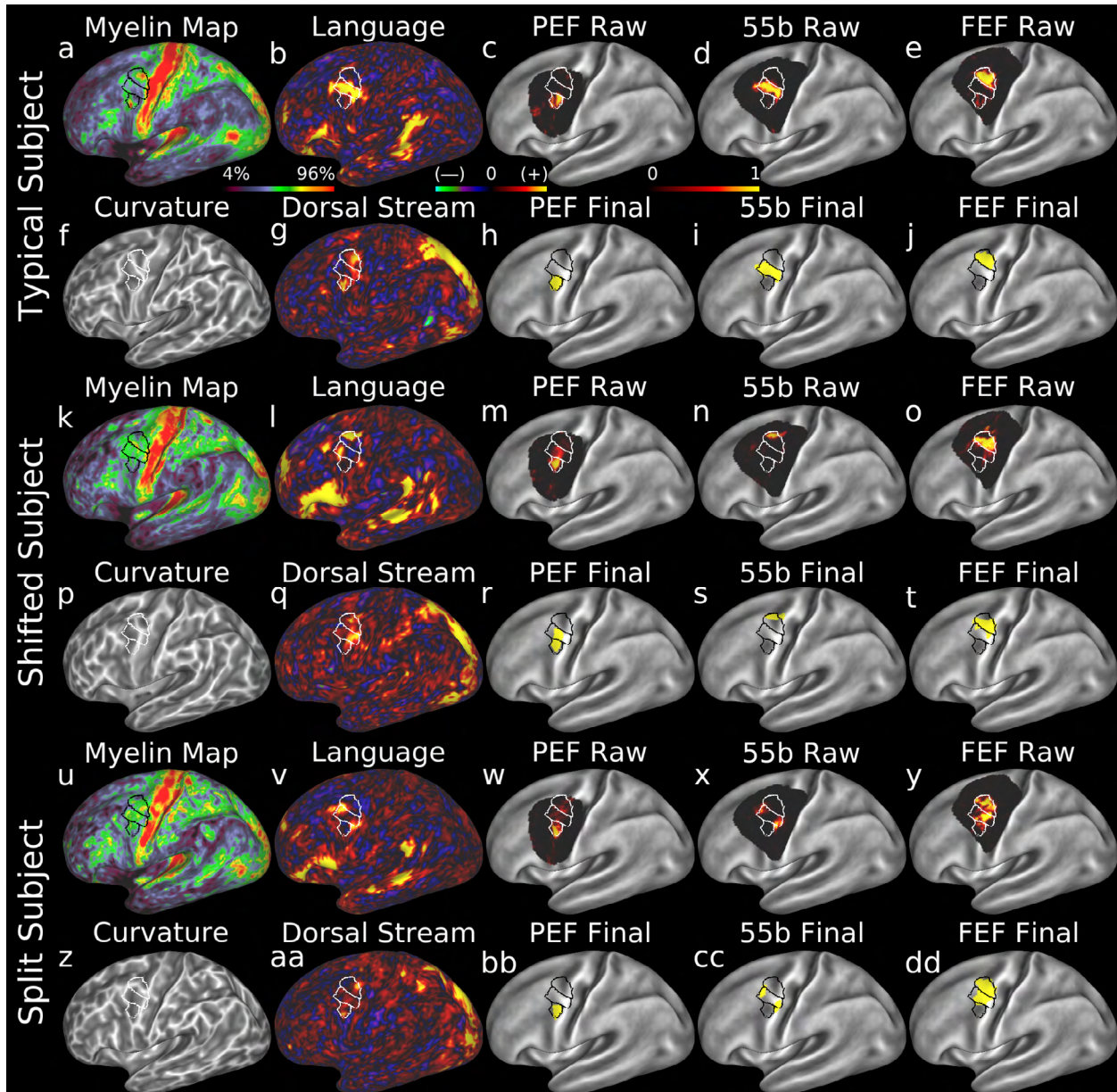


**Figure 8 | Atypical splits of area 55b.** Here we show examples of a second type of topological incompatibility between the typical layout of areal features near area 55b and a minority of subjects with alternative ‘split’ layouts. Row 1 shows the group average areal features: **a**, a  $d=40$  RSN map corresponding to parts of the language network, **b**, a myelin map, **c**, a LANGUAGE Story-Baseline task contrast beta map, **d**, a  $d=40$  RSN map corresponding to the dorsal visual stream, and, **e**, a mean curvature map illustrating the folding pattern. The group delineation of areas 55b, FEF, and PEF are white or black lines. **f–j** show these maps in a typical individual subject. **k–o** show an atypical subject whose area 55b is split and FEF and PEF are joined. **p–t** show a second atypical subject with a similar pattern. In both cases the split of 55b is visible in multiple independent areal features. Data at <http://balsa.wustl.edu/Wrn2>.

#### 1.4 Automated individual subject multi-modal parcellations

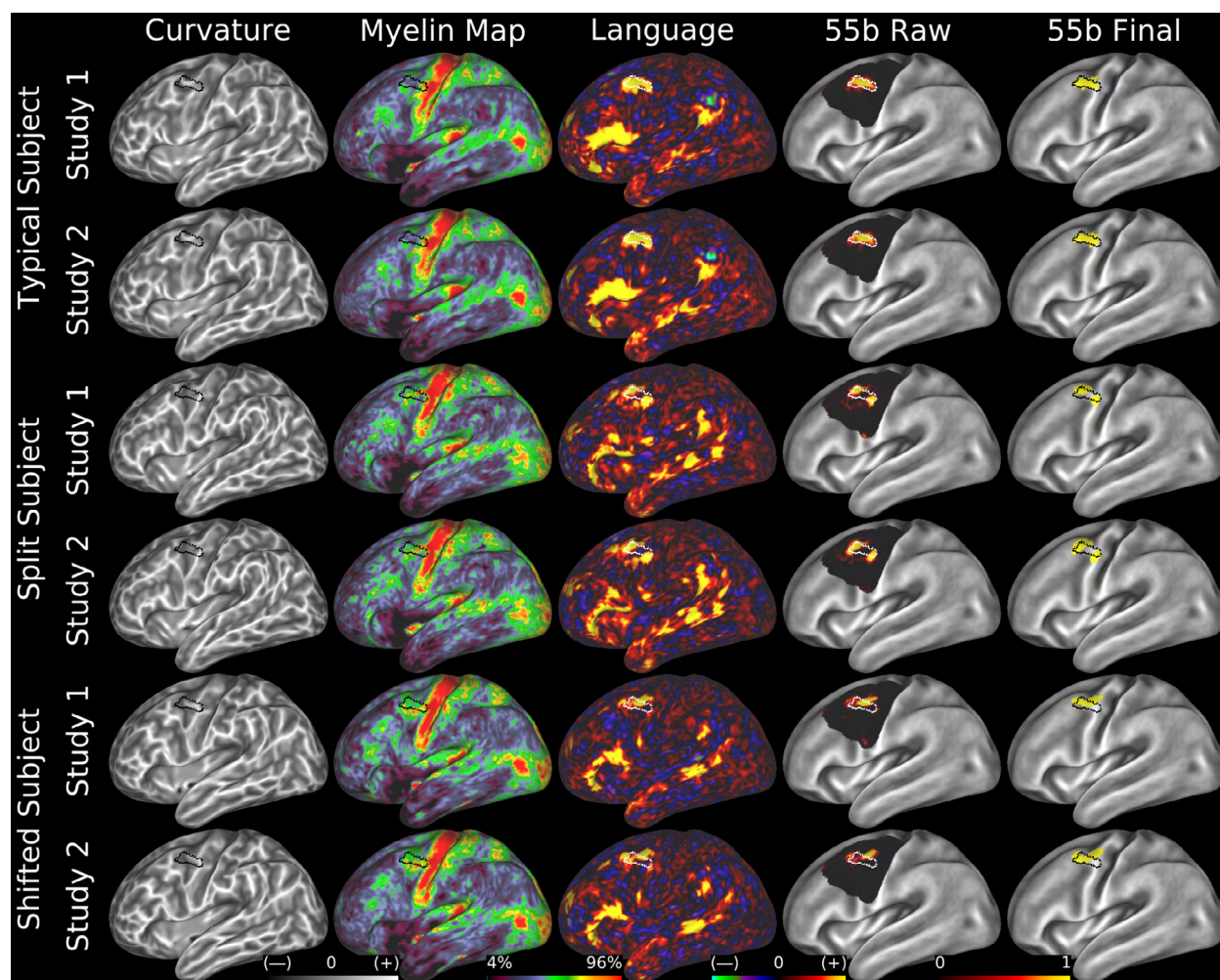
The differing areal topologies of 55b and its neighbors not only demonstrate the need for individual subject parcellation (as opposed to atlas-based parcellation) but also serve as an excellent test case to evaluate critically whether an individual subject parcellation method is performing as desired. Figure 9 shows parcellation results for three example subjects, with Fig. 9a–j showing a typical 55b subject, Fig. 9k–t showing a shifted 55b subject, and Fig. 9u–dd showing a split 55b subject. The classifier output from the typical subject shows areas PEF (column 3), 55b (column 4), and FEF (column 5) in roughly the same places as in the group average (i.e. this subject has been well aligned with the areal-feature-based registration). In the shifted 55b subject, FEF and 55b have swapped places, and PEF extends into the region normally classified as 55b. Despite the atypical topology, the classifier was able to identify all three areas in their atypical locations based on their multi-modal areal fingerprints. In the split 55b subject, the classifier indicates that 55b is divided into two pieces, whereas FEF and PEF adjoin one another. Examples like this motivated us not to enforce a strict spatial contiguity criterion on the individual subject

parcellations. Instead, these split-area cases argue that it is possible for some areas to be spatially non-contiguous in some individuals, even if they are contiguous in most individuals. That the classifier was able to identify these atypical cases indicates that it is performing as desired, identifying cortical areas in many individuals even when their areal topology differs from the group average.



**Figure 9 | Areal classification of typical and atypical areas 55b.** Here we show classification of areas PEF, 55b, and FEF in three example individual subjects with typical, shifted, and split 55b areal topologies respectively (group parcels are black or white outlines). **a, k, & u** show the individual subject myelin maps. **b, l, & v** show the individual subject language resting-state network maps ( $d=40$ ). **f, p, & z** show the individual subject folding maps. **g, q, aa** show the individual subject dorsal stream visual network maps ( $d=40$ ). **c, d, e, m, n, o, w, x, & y** show the raw classifier output likelihood maps (scaled 0-1 black to yellow) for PEF, 55b, and FEF from left to right. **h, i, j, r, s, t, bb, cc, & dd** show the final parcels after combining across classifiers for PEF, 55b, and FEF. For all subjects, the areal classifier is able to identify area 55b sensibly. Data at <http://balsa.wustl.edu/QxB9>.

Another key indicator of the robustness of our classification method (and of the quality of the HCP data and analysis) comes from evaluating the reproducibility of results across repeat scans of individual HCP test-retest subjects. Figure 10 shows results from three subjects who were scanned through the full HCP protocol twice and processed through completely separate runs of the HCP's processing pipelines. Rows 1 and 2 show test-retest results for a subject with a typical 55b topology, rows 3 and 4 show results for a split-topology subject, and rows 5 and 6 show results for a shifted-topology subject. Columns 1 – 3 indicate a high degree of reproducibility for cortical folding (column 1, correlations for cases 1/2/3 are  $r=0.88/0.84/0.83$ ), cortical myelin (column 2, correlations for cases 1/2/3 are  $r=0.91/0.88/0.87$ ), and RSN maps (column 3, correlations for the illustrated  $d=40$  RSN maps for cases 1/2/3 are  $r=0.64/0.62/0.5$ ). As shown in columns 4 and 5, the classifier was able to delineate and identify area 55b in a consistent location in all three cases (correlation coefficients for the final parcels shown in cases 1/2/3 are  $r=0.87/0.79/0.82$ ). Thus, while not as high as the group map correlations, the individual subject maps are reasonably reproducible.



**Figure 10 | Reproducible areal classification of typical and atypical areas 55b.** Here we show individual subject curvature maps (column 1), myelin maps (column 2),  $d=40$  RSN maps of the language network (column 3), classifier raw probability maps (column 4), and final parcel maps (column 5) of subjects scanned and analyzed through the HCP protocol and software pipelines twice. The group 55b definition is shown in a

black or white outline. Rows 1 and 2 show a subject with a typical 55b topology, rows 3 and 4 show a split 55b topology, and rows 5 and 6 show a shifted topology. Data at <http://balsa.wustl.edu/WPPn>.

Figure 11 shows the group parcellation and compares it with parcellations from a different set of three HCP test-retest subjects. The first row shows the group parcellation, which serves as the reference for rows 2-6 (reproducibility scores are correlation or Dice coefficients of the parcellations represented as 180 binary ROIs). Results are shown in row 2 for the subject with the highest overall parcellation reproducibility ( $r=0.79$ , Dice=0.75), row 3 for the subject with the median overall parcellation reproducibility ( $r=0.77$ , Dice=0.72), and row 4 for the subject with the lowest overall parcellation reproducibility ( $r=0.73$ , Dice=0.65), with individual-subject parcel boundaries shown in blue (test), red (retest), and purple (overlap) and group-average boundaries overlaid in black. Inspection of these maps reveals numerous examples in which deviation of an individual area boundary from the group mean is similar in both test and retest data but also examples of differences (likely errors). Rows 5 and 6 show the mean reproducibility of each area across the 27 test-retest subjects for the correlation and Dice measures. Larger areas tend to have higher reproducibility than smaller areas. While not as high as the group MPM reproducibility, the individual MPMs are reasonably reproducible. A topic for future study will be elucidating what causes differences across runs (see discussion below, #2.2-2.3).

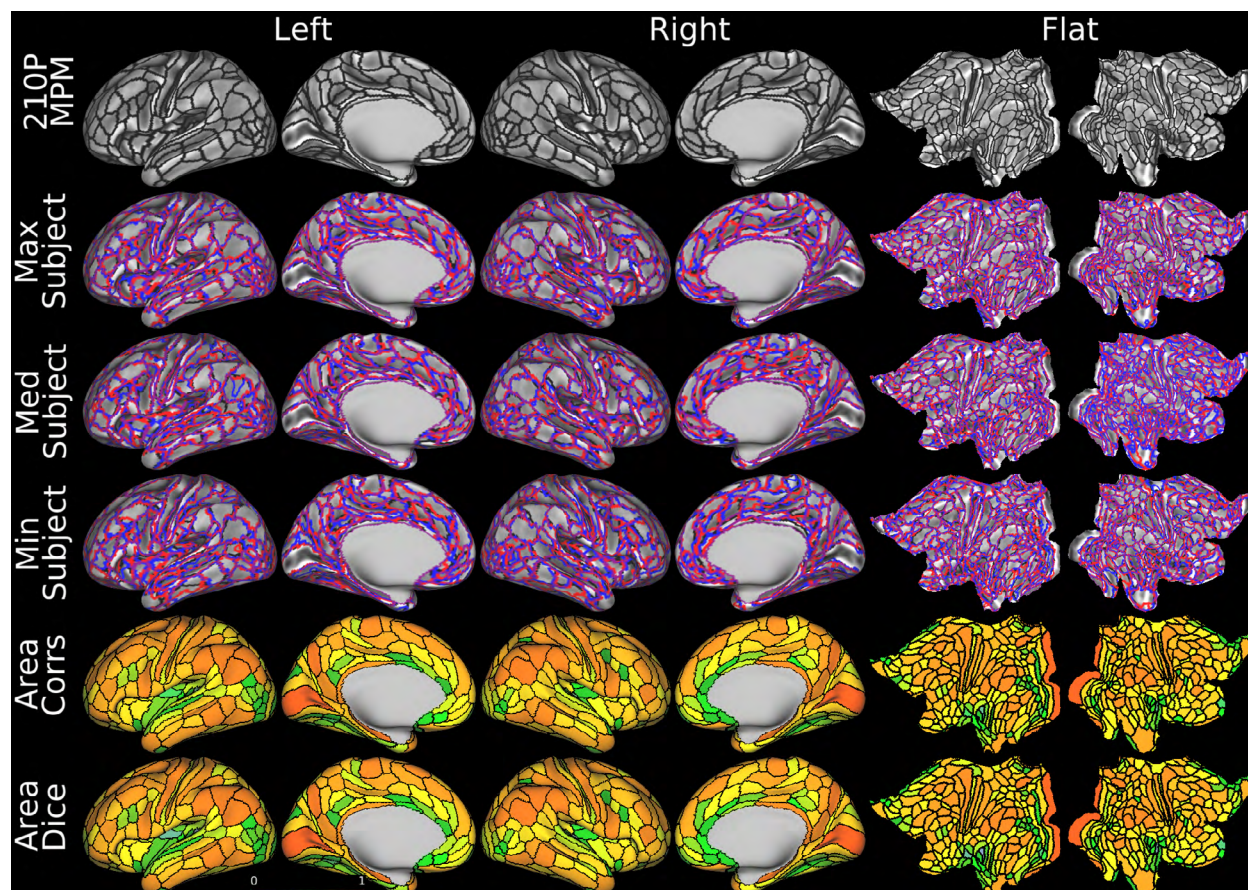


Figure 11 | **Individual subject parcellation reproducibility.** Here we show the group parcellation in comparison with three individuals' parcellations (who were each scanned twice). Row 1 shows the group parcellation (black outlines). Row 2 is the individual having the highest parcellation reproducibility (blue

outlines are 'test' individual areal boundaries and red outlines are 'retest' areal boundaries with purple as overlap and group boundaries as an underlay of black outlines). Row 3 is a second individual, who had the median parcellation reproducibility. Row 4 is a third individual, who had the lowest parcellation reproducibility. Rows 5 and 6 show the mean parcel reproducibility for each area across all 27 subjects that were scanned and analyzed twice for both the correlation and Dice measures (the full range is  $r=0.93$ ,  $r=0.44$ ,  $\text{Dice}=0.93$ ,  $\text{Dice}=0.41$ ). Data at <http://balsa.wustl.edu/W72X>.

### 1.5 Variability in surface areas across group cortical areas and in individual subjects

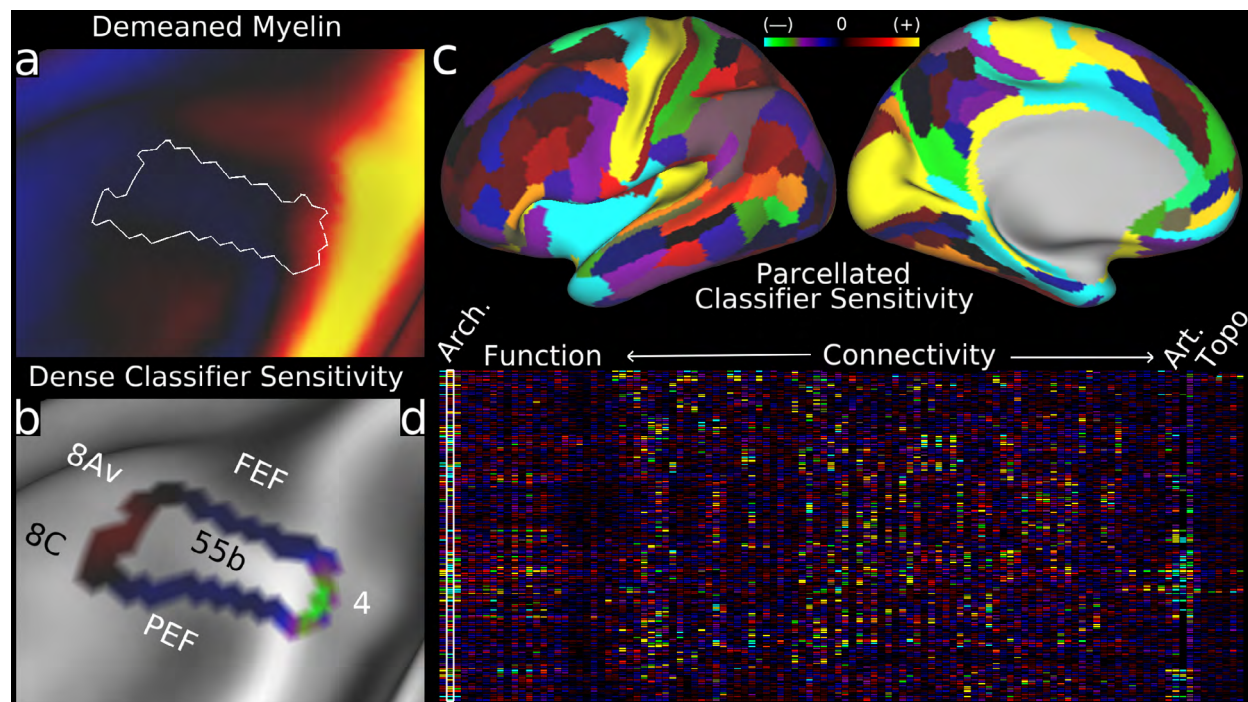
The availability of individual-subject parcellations in which a large majority of areas were identified in a large majority of individuals enabled interesting initial analyses of the sizes of cortical areas in the group average and variability in areal sizes across individuals. In individuals from the 210P group, the average size (i.e. surface area) of each cortical parcel spans a 26-fold range. V1 is the largest area (3292 mm<sup>2</sup> and 3229 mm<sup>2</sup> in the left and right hemispheres); nearly 80% of areas (142/180) are substantially smaller and within a four-fold range (150 – 600 mm<sup>2</sup>, averaged across hemispheres). Across individuals, V1 is the least variable, with a total range slightly exceeding 2-fold, and a full-width at half maximum (FWHM) of 32% of the mean on the left and 30% on the right. Other areas larger than 500 mm<sup>2</sup> have variability with FWHM between 40% and 80% of the mean, while areas smaller than 500 mm<sup>2</sup> are generally more variable (mostly with FWHM 50% - 120% of the mean, in all cases excluding areas that were not identified in a given individual). These values represent the upper range of possible 'true' anatomical variability, given known methodological variability from subjects who had repeat scanning sessions. These data are amenable to future analyses of how individual variability in areal sizes correlates with behavioral and other phenotypic characteristics.

### 1.6 Peeking inside the areal classifier

We can "peek inside the black box" to examine the areal fingerprints learned by the classifier and ascertain to what degree it used the multiple modalities. We measured how sensitive the classifier is to changes in each feature map for each area as an indicator of the influence that each feature map has on generating the classification of each area (see Methods section on the cortical areal classifier). This sensitivity measure is positive or negative depending on whether the feature's value is larger or smaller inside the area, and its absolute magnitude indicates how important the feature is to the area's classification relative to its neighbors.

This sensitivity measure can be interrogated in multiple ways, as illustrated in Figure 12. Fig. 12a-b show the sensitivity measure at a 'dense' (vertex-wise) level for the demeaned myelin map in and around area 55b. The classifier 'sees' (Fig. 12b) area 55b as having much less myelin than area 4 posteriorly, slightly less myelin than FEF (superiorly) and PEF (inferiorly), and slightly more myelin than areas 8Av and 8C anteriorly, consistent with the pattern in the myelin map (Fig. 12a). We summarized this sensitivity measure for each area across all of its borders with neighboring areas by taking the maximum absolute value while retaining the sign, as shown in the parcellated sensitivity map for the myelin feature in Fig. 12c. Fig. 12d shows the matrix of these values for all 180 areas in the left hemisphere and all 112 features. Overall, all modalities are indeed useful for delineating and identifying cortical areas with the fully automated approach, just as with the semi-

automated neuroanatomical approach. Of the different areal properties, architectural features and resting state connectivity features drove the classifier's choices most strongly. Resting state visuotopic features were heavily used for specific areas (particularly early visual areas), but as expected are of limited value elsewhere. Task fMRI is also used, but clearly not as much as the other modalities (it has few yellows and cyans indicating high positive or negative sensitivity values). One possible explanation is that task fMRI maps are noisier than other feature maps and the information they do contain is largely redundant with other features (e.g. resting state) that have higher contrast to noise ratios in individual subjects.



**Figure 12 | Signed classifier sensitivity measures.** **a**, shows the demeaned group average myelin map (what the classifier operates on) with area 55b outlined in white. **b**, shows the sensitivity measure for the myelin map and area 55b (only measured near the 55b border). Note that the measure is especially negative for the 55b/4 border (posteriorly), modestly negative for the 55b/FEF and 55b/PEF borders (superiorly and inferiorly), and modestly positive for the anterior border of 55b with 8Av and 8C. **c**, shows the maximum absolute value (shown with sign) of each area for the myelin map feature, indicating that myelin is useful for defining many, but by no means all, cortical areas. **d**, shows the matrix of all areas and all features (myelin is outlined white). The feature categories Architecture (Arch.), Function, Connectivity, Artifacts (Art.), and Topography (Topo.) are labeled along the top of the matrix. Data at <http://balsa.wustl.edu/W60I>.

## 2. Supplementary Discussion

### 2.1 Group average cortical parcellation.

The 180 cortical areas per hemisphere identified in this study are likely fewer than the actual number, because: (i) we have focused on particularly robust borders (consistent across at least two measures with a Cohen's effect size of at least 1 for 96% of borders), so it is unlikely that parcels were split too finely; and (ii) some parcels likely could be split

more finely, as our parcellation may be too coarse in some regions. For example, in orbitofrontal cortex we were unable to replicate known architectonic parcels (Ongur et al., 2003), likely in part because of susceptibility induced fMRI signal loss in that region. We also did not identify as many retinotopic areas in higher visual cortex as some studies have reported (Abdollahi et al., 2014; Wandell and Winawer, 2015; Wang et al., 2005). We identified some parcels as ‘complexes’ because we could not identify convincing subdivisions of a region that was previously reported to have a finer-grained parcellation. Hence, 180 areas is likely a lower bound on the actual number of areas, and ~200 per hemisphere might be closer to reality (though of course there will be finer levels of organization than cortical areas; see also point (v) in #2.8 below).

In this context, we hope that our parcellation, which we identify as HCP-MMP1.0 (Human Connectome Project Multi-Modal Parcellation v1.0), serves as version 1.0 of a progressively improving entity (analogous to open source software in which bugs may be fixed and new features implemented). Refinements to our cortical parcellation may come from multiple sources: (i) additional HCP data from the final 3T release (~1100 subjects with MRI data to be released in ~June, 2016) and/or the 7T data (~180 subjects and including retinotopy and movie paradigms, as well as rfMRI and dMRI). (ii) HCP tractography data, which may provide evidence for finer-grained parcellation in some regions (though see below #2.5); (iii) higher resolution structural and fMRI data that enable laminar-based cortical parcellation of myelin maps and resting-state networks; (iv) customized task fMRI experiments with high quality spatial localization that focus on particular regions; and (v) improved data acquisition, such as whole brain spin echo fMRI at high field that would ameliorate regional gradient echo fMRI signal loss. It will be important to use areal-feature-based registration and de-drifting in future studies to maximize the accuracy of cross-study comparisons, including with this parcellation.

The bilateral symmetry in the position, size, and functional connectivity profile of corresponding areas in the two hemispheres is striking (see Main Text Figure 3). The pattern is much more symmetric than those reported in several recent fully automated parcellations based only on resting-state functional connectivity (Gordon et al., 2014; Shen et al., 2013; Yeo et al., 2011), but is consistent with the bilateral symmetry generally reported in observer-independent postmortem architectonic parcellations (e.g. (Caspers et al., 2013)). Nevertheless, a number of interesting asymmetries are evident, some of which are described in the Supplementary Neuroanatomical Results.

## *2.2 Individual-subject parcellations.*

Using a machine learning classifier, we successfully identified 98% of areas in the 210P group and 96.6% of areas in the 210V group that was not used to generate the parcellation. An important feature of the areal classifier is its ability to delineate and identify the ‘non-standard’ topology of some areas in some individuals. Area 55b was a notable hotspot for nonstandard topologies (indicated by higher average registration distortion in this region), including shifts in the local order of areas and splitting of area 55b into topologically discontinuous segments separated by joined areas FEF and PEF. Topological irregularities are evident for other areas as well, but their overall pattern and frequency, as well as the heritability of such patterns have yet to be systematically examined. It also is intriguing to speculate on how topological irregularities arise during



development, what degree of heritability they show (e.g., to what degree are atypical patterns shared by monozygotic twins), and whether they are associated with behavioral or other phenotypic correlates, but such topics are outside the scope of this study.

Although the areal classifier detects areas at a high rate, it does not detect them at a rate of 100% across all subjects. There are several possible causes for areas not detected by the classifier: 1) Some areas may neurobiologically truly not exist in some subjects. 2) Noise or artifacts (e.g. gradient echo fMRI dropout or the effects from large veins) may make some areas undetectable in some subjects (particularly smaller areas, areas near dropout regions, and areas that are not as distinct from their neighbors as are better detected areas). 3) Nonstationarity in the data (particularly in rfMRI scans, see #2.3) may lead to some areas being undetectable (if they are sufficiently similar to a neighbor across the available data to not be separated by the areal classifier). Better understanding the causes of undetected areas is an issue for future work: with more or better data, these areas might be detectable, or their absence might persist, suggesting that they represent genuine inter-subject variability like the atypical areal topologies. For now, we recommend that these areas be treated as having zero surface area for morphometric analyses and be controlled for using covariates of no interest for functional/connectivity cross-subject designs (i.e. a regressor with all zeros except for a 1 for the missing area in each subject that is missing it). Also, the group parcellation can be used as an atlas to parcellate (i.e. average data within the parcels) individual-subject data based on the individual subject to atlas alignment of the areal feature-based registration, though at the cost of reduced accuracy for atypical topologies that cannot be aligned using a topology preserving registration.

### *2.3 Aligning and parcellating future 'HCP-style' datasets.*

In addition to high resolution T1w and T2w, the HCP acquired one hour of resting state fMRI data and one hour of task fMRI data (total across all tasks), which were used to derive the original semi-automated parcellation and for areal classification. MSMAll areal-feature-based alignment was driven solely by the resting state fMRI data and the T1w/T2w myelin maps. We tested the areal classification without the task fMRI data, finding that it detected areas at almost the same rate (96.4% vs 96.6% for the 210V data). Future work is needed to determine the minimum amount and type of fMRI data required for successful application of MSMAll and the areal classifier, but we can make several reasonable predictions: 1) It is likely that what will prove most important is the total amount of available fMRI data that has been cleaned using data-driven methods such as ICA+FIX (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014) and then combined across runs. Whether the data is resting state fMRI, task fMRI, or movie fMRI is likely to be less important (for example, HCP resting state and task fMRI data, when analyzed similarly, are reported to produce similar results, (Cole et al., 2014)). 2) The minimum amount of fMRI data needed to produce accurate MSMAll registrations and areal classifications is not yet known. Two considerations drive the need for more fMRI data. The first is that more fMRI data will improve the quality of connectivity estimates relative to unstructured (i.e. Gaussian) noise. The second is that fMRI data contain substantial nonstationarities, and thus connectivity estimates from a limited amount of time may not reflect the true average connectivity estimates, (irrespective of the effects of unstructured noise). This is expected

because the connectivity is thought to reflect what is going on in a person's brain at a given time, and shorter time windows are more likely to be different from a large average time window. The first consideration is amenable to improvement using better scanners (e.g. higher field strength), better head coils, and better pulse sequences (spin echo might help with dropout or large veins), but the second will be an intrinsic limitation of this kind of functional imaging, requiring that enough minutes of fMRI data are acquired. Perhaps this issue could be improved by using a more controlled experimental setting (e.g. a naturalistic movie that both holds the subject's attention and covers a wide range of stimuli). In general, optimizing the quality of the individual subject multi-modal parcellations (increasing the areal detection rate and the test-retest reproducibility of the parcellations) is an important subject for follow up work, enabled by what has been presented here.

Altogether, we anticipate that the methods described here will work for the 'HCP-style' data acquisition recommendations being used in forthcoming NIH-funded projects on the Lifespan Connectome and on Connectomes Related to Human Disease even though these projects will acquire less data per subject (see also Glasser et al., 2016). This will likely include ~30 min of multiband rfMRI data and another ~20 min of multiband tfMRI data. These datasets can likely be cleaned and combined to enable identification of RSNs using close to the 1 hour of rfMRI data used in this study for MSMAll alignment and the areal classifier. Other topics for future work include investigating how poor subject compliance (subject motion, subject going to sleep in the scanner, etc.) affects these methods. We strongly recommend the use of a method such as ICA+FIX (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014) in order to remove spatially specific structured noise. The methods presented here for alignment and individual subject parcellation will not otherwise be affected by global signal differences (because they rely on multiple regression to generate resting state network maps, which represent partial betas more akin to partial correlation than full correlation). Thus, decisions to use or not use global signal regression or partial correlation can be postponed until one has generated parcellated timeseries. Ideally, however, one would eliminate the artifactual, likely physiologically induced (Golestani et al., 2015) global noise while retaining the genuine global neural signal and hopefully move the field past the divisive debate about global signal regression (see Glasser et al., 2016). Low spatial (>2.6 mm, the mean thickness of cortex) and temporal resolution (>1 s TR) data may not give as good results as the 2 mm, 0.72 s data used in this study because of blurring across cortical folds and worse data cleanup with slow TR data (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014).

#### *2.4 Circularity Considerations.*

One potential concern regarding areal-feature-based alignment is whether it introduces circularity into the analyses. Two points are worth raising in this context: 1) Previous analysis of folding patterns, cortical thickness, or volume-based morphometric analyses are susceptible to a similar criticism (given that folding-based surface registration or T1w image intensities in a volume registration are used to drive the alignment in these cases). 2) We previously showed that areal-feature-based alignment using myelin maps and resting state network maps produces higher cross-subject statistics for an independent areal feature (task fMRI) that was not used to drive the registration, relative to alignment using folding-based registration (Robinson et al., 2014). Also, our implementation of the

MSMAll areal feature-based registration includes relatively strong regularization, generating substantially less distortion than standard folding-based methods (e.g. FreeSurfer, see Methods section on image preprocessing), and making it less likely that the registration would convert random noise into imputed areal features.

### *2.5 Diffusion imaging and tractography.*

For technical reasons, we did not use diffusion imaging and tractography data as a source of information for our parcellation, even though tractography has been used in other studies to distinguish between gray matter parcels based on ‘connectional contrast’ (Behrens and Johansen-Berg, 2005; Jbabdi and Behrens, 2013). A primary reason was that tractography is susceptible to folding-related biases, including a tendency for streamlines to terminate preferentially on gyral crowns and to avoid sulcal fundi (Reveley et al., 2015; Van Essen et al., 2014). Given such biases, gradients in tractography connectivity may tend to correlate more with folding patterns for artifactual rather than neurobiological reasons. Understanding and improving the performance of tractography, especially near the gray/white border as fibers exit and enter the gray matter, is a topic of on-going work (e.g., (Canales-Rodríguez et al., 2015)). The ability to confirm consensus multi-modal borders using more advanced tractography algorithms will be a good test of how well these algorithms perform in individuals and group averages. One automated way of doing this would be to include tractography features from the 210P and 29T datasets in a classifier training and see if the tractography behaved more like the resting state connectivity or more like the task fMRI in an analysis like Figure 12 (i.e. does the areal classifier make use of the tractography?). Should this be successful, tractography may aid in identifying new boundaries as well. In any event, the current HCP parcellation provides a neuroanatomical framework that will enable quantitative comparisons of parcellated connectomes based on tractography with those based on functional connectivity, at the level of individual subjects as well as group averages.

### *2.6 Integration with other parcellations.*

Substantial work remains in order to integrate our parcellation with information currently only available from post mortem histological methods, such as cytoarchitecture, laminar myeloarchitecture, and transmitter-receptor architecture (e.g., (Caspers et al., 2013; Zilles et al., 2015)). Simply bringing histological data into register with MRI data and into traditional MRI standard spaces has been a highly challenging but very successful methodological undertaking (Caspers et al., 2013). That said, it will be critical to apply the same sorts of spatial localization approaches to such histological data as was used for the in vivo data in this study to enable precise neuroanatomical comparisons between post-mortem and in vivo data. Early steps in this direction (Fischl et al., 2008) proved beneficial both in the present study and in previous work (Glasser and Van Essen, 2011).

### *2.7 Subcortical and cerebellar parcellation.*

Many of the methods and approaches used in the present study could be readily extended to subcortical gray matter. This will be of particular interest for the HCP 7T data,

as the fMRI data have markedly higher CNR at 7T relative to the 3T scans and higher resolution at 1.6 mm vs 2 mm. Future higher resolution versions of the CIFTI grayordinates space may enable parcellation of additional subcortical structures (e.g. the claustrum).

It will also be of interest to use HCP data to parcellate the cerebellum, along the lines of a previous RSN-based cerebellar segmentation (Buckner et al., 2011), but using the rfMRI and tfMRI data from the HCP 3T and/or 7T scans. Ideally, such analyses would be done using cerebellar surface reconstructions of individual subjects, and then mapped to a group average cerebellar atlas surface. However, because cerebellar cortex is much thinner than cerebral neocortex, this is not yet feasible using existing segmentation algorithms, even though the high-resolution (0.7 mm) HCP T1w and T2w scans reveal many details of cerebellar folding. Nonetheless, a volume-based HCP cerebellar parcellation followed by mapping to an existing cerebellar surface-based atlas (Van Essen, 2002) would represent a considerable advance.

### *2.8 Sub-parcels and internal areal heterogeneity.*

In general, a brain parcel is a region that has greater commonality of features within the parcel than with neighboring parcels (and this is explicitly what the areal classifier learns as a multi-modal areal fingerprint). However, within-parcel heterogeneity is profoundly important for understanding detailed aspects of brain circuitry and function. Internal heterogeneity extends across multiple levels, including within-area modularity and columnar organization all the way down to cellular and subcellular levels, but the heterogeneity accessible to the noninvasive 3T HCP neuroimaging data studied here is much coarser. We identified several types of within-parcel heterogeneity. (i) The somatosensory-motor strip contains five well-defined sub-regions, representing the lower limb, trunk, upper limb, face, and (surprisingly) a small patch of motor cortex likely representing muscles that control squinting of eyes (Meier et al., 2008), among other things. (ii) Within visual cortex, we demonstrated orderly gradients in the visuotopic maps of polar angle and eccentricity, which were made using resting state data. Our results are analogous to previous reports of visuotopy based on resting state fMRI (Gravel et al., 2014; Heinzle et al., 2011) but they benefitted from a number of analysis methods introduced here (see Supplementary Methods). (iii) Analogous orderly topographic gradients were evident in several higher cognitive regions – e.g., language-related areas 55b, PSL, and SFL (see Supplementary Neuroanatomical Results #15 Figure 18) though their functional significance is currently unknown. (iv) Many areas in our parcellation showed internal heterogeneity in one or more features, including task-fMRI, which were modest in magnitude and/or not consistent across multiple modalities. The reproducibility across independent datasets (see Figures 1-5 above) suggests that these heterogeneities mostly represent neurobiologically interesting variation, though some may reflect task-induced ‘artifacts’ (e.g. button box pressing, activation of only part of the visual field, etc). (v) Finally, although the topographic organization of early visual, somatosensory-motor, and auditory regions is reasonably well understood, we know much less about such organization in higher cognitive regions. One major impediment to this understanding in humans has been the use of brain imaging analysis methods that blur the fine details of functional and structural organization, obscuring such topographic patterns. Another is

that many of the higher cognitive regions in question have expanded dramatically in humans relative to the intensively and invasively studied macaque monkey (Van Essen and Dierker, 2007). These regions also have dramatically differing patterns of structural connectivity (Rilling et al., 2008), and thus are likely very different from putatively homologous regions in monkeys. Because the topographic organization of early somatosensory-motor cortex is well understood, we were able to separate which gradients defined topographically organized cortical areas (4, 3a, 3b, 1, and 2) from topographic subregions (F, E, UL, T, LL) as described in Supplementary Neuroanatomical Results #6. However, in higher cognitive regions, we don't have the benefit of such prior knowledge. Thus, future versions of the human multi-modal parcellation may be revised to classify some of the subdivisions reported here as subareas rather than areas as new information becomes available and the topographic organization of human cerebral cortex becomes better understood.

Additional topics of interest include analyses of the patterns of interhemispheric symmetry and asymmetry, and specific aspects of functional specialization (e.g., degree of segregation vs partial overlap for regions activated by faces, body parts, tools, and places). Some of these are covered in the Supplementary Neuroanatomical Results. Others can be examined in detail by downloading and analyzing the datasets associated with this study (see below #2.10).

### *2.9 Analysis strategies and cross study comparisons.*

We have made significant efforts to enable other investigators to compare their data as precisely as possible with our parcellation through extensive releases of software tools that enable dramatically improved spatial localization: the CIFTI grayordinates neuroimaging analysis framework (Glasser et al., 2013), areal-feature-based surface registration with MSM (Robinson et al., 2014), group average registration drift removal during template generation (Abdollahi et al., 2014), and an extensive datasharing effort described below.

In general, neuroimaging studies that are mainly interested in what brain areas are activated/connected/different in some way will likely benefit from parcellated analyses due to their increased statistical sensitivity and power. Additionally, the trained classifier will hopefully enable identification of these cortical areas in individuals from future studies imaged with high quality "HCP-style" data acquisition and analysis strategies, such as the forthcoming Lifespan Connectome or Connectomes Related to Human Disease projects, which meet the minimal requirements of the HCP Preprocessing Pipelines, the MSMAll registration pipeline, and the individual subject areal classifier (i.e. high spatial and temporal resolution T1w, T2w, fMRI, and spin echo b0 fieldmap (Glasser et al., 2013; 2016)).

### *2.10 Sharing data via the BALS database.*

The ConnectomeDB database (Hodge et al., 2015; Marcus et al., 2013); <https://db.humanconnectome.org> provides a user-friendly source for accessing unprocessed, minimally preprocessed, and subject-wise extensively analyzed HCP datasets. However, ConnectomeDB is not designed for sharing group average data or results from

neuroimaging studies such as this study. For this purpose, we designed and recently implemented the BALSAs (Brain Analysis Library of Spatial maps and Atlases) database (Van Essen et al., 2016) (<http://balsa.wustl.edu>). BALSAs is organized around scene files that provide all of the information needed to upload, download, and visualize in Connectome Workbench the neuroanatomical illustrations associated with this study. In essence, each scene in a scene file includes detailed information about the files used to generate the display of each of the figure panels. This includes the specific geometric configurations, overlays of different data types, labels and annotations used in each figure panel (generally, a separate 'tab' in the 'tile tabs' display enabled in `wb_view`, the Connectome Workbench visualization software platform). The process of uploading a scene file to BALSAs imports all of the data files needed to regenerate its constituent scenes and includes the directory structures needed for successful scene visualization upon downloading. In addition to providing the scenes for each figure in this study, there will be additional reference data related to this study available in BALSAs in the future.

Users accessing BALSAs can navigate to the available datasets associated with the current study and inspect previews of each scene in each of the available scene files (one for each of the Main Text (<https://balsa.wustl.edu/sceneFile/show/ILMz>), Supplementary Results and Discussion (<https://balsa.wustl.edu/sceneFile/show/X59l>), Supplementary Methods (<https://balsa.wustl.edu/sceneFile/show/PrGD>), and Supplementary Neuroanatomical Results (<https://balsa.wustl.edu/sceneFile/show/D4zL>)). Also, we have provided a Study Dataset (<https://balsa.wustl.edu/sceneFile/show/L731>) that includes the group average data and gradients used to generate the parcellation, the original semi-automated parcellation, the 210P and 210V probabilistic areas, and the 210P and 210V group MPMs. Once analysis is completed on the full ~1100 subject HCP dataset, a Reference Dataset will be provided on the BALSAs database. Users can download the data after obtaining a BALSAs user account and agreeing electronically to HCP Open Access Data Use Terms.

## References

- Abdollahi, R.O., Kolster, H., Glasser, M.F., Robinson, E.C., Coalson, T.S., Dierker, D., Jenkinson, M., Van Essen, D.C., Orban, G.A., 2014. Correspondences between retinotopic areas and myelin maps in human visual cortex. *Neuroimage* 99, 509-524.
- Behrens, T.E., Johansen-Berg, H., 2005. Relating connectional architecture to grey matter function using diffusion imaging. *Philos Trans R Soc Lond B Biol Sci* 360, 903-911.
- Buckner, R.L., Krienen, F.M., Castellanos, A., Diaz, J.C., Yeo, B.T., 2011. The organization of the human cerebellum estimated by intrinsic functional connectivity. *Journal of neurophysiology* 106, 2322-2345.
- Canales-Rodríguez, E.J., Daducci, A., Sotiropoulos, S.N., Caruyer, E., Aja-Fernández, S., Radua, J., Mendizabal, J.M.Y., Iturria-Medina, Y., Melie-García, L., Alemán-Gómez, Y., 2015. Spherical Deconvolution of Multichannel Diffusion MRI Data with Non-Gaussian Noise Models and Spatial Regularization. *PloS one* 10, e0138910.
- Caspers, S., Eickhoff, S.B., Zilles, K., Amunts, K., 2013. Microstructural grey matter parcellation and its relevance for connectome analyses. *Neuroimage* 80, 18-26.
- Cole, M.W., Bassett, D.S., Power, J.D., Braver, T.S., Petersen, S.E., 2014. Intrinsic and task-evoked network architectures of the human brain. *Neuron* 83, 238-251.

- Fischl, B., Rajendran, N., Busa, E., Augustinack, J., Hinds, O., Yeo, B.T., Mohlberg, H., Amunts, K., Zilles, K., 2008. Cortical folding patterns and predicting cytoarchitecture. *Cereb Cortex* 18, 1973-1980.
- Glasser, M.F., Goyal, M.S., Preuss, T.M., Raichle, M.E., Van Essen, D.C., 2014. Trends and properties of human cerebral cortex: correlations with cortical myelin content. *Neuroimage* 93 Pt 2, 165-175.
- Glasser, M.F., Smith, S.M., Marcus, D.S., Andersson, J. Auerbach, E.J., Behrens, T.E.J., Coalson, T.S., Harms, M.P., Jenkinson, M. Moeller, S., Robinson, E.C., Sotiropoulos, S.N., Xu, J., Yacoub, E., Ugurbil, K., and Van Essen, D.C., 2016. The Human Connectome Project's neuroimaging approach. *Nature Neuroscience* (in press).
- Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., Van Essen, D.C., Jenkinson, M., Consortium, W.U.-M.H., 2013. The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage* 80, 105-124.
- Glasser, M.F., Van Essen, D.C., 2011. Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. *J Neurosci* 31, 11597-11616.
- Golestani, A.M., Chang, C., Kwinta, J.B., Khatamian, Y.B., Jean Chen, J., 2015. Mapping the end-tidal CO<sub>2</sub> response function in the resting-state BOLD fMRI signal: spatial specificity, test-retest reliability and effect of fMRI sampling rate. *Neuroimage* 104, 266-277.
- Gordon, E.M., Laumann, T.O., Adeyemo, B., Huckins, J.F., Kelley, W.M., Petersen, S.E., 2014. Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. *Cereb Cortex*.
- Gravel, N., Harvey, B., Nordhjem, B., Haak, K.V., Dumoulin, S.O., Renken, R., Ćurčić-Blake, B., Cornelissen, F.W., 2014. Cortical connective field estimates from resting state fMRI activity. *Frontiers in neuroscience* 8.
- Griffanti, L., Salimi-Khorshidi, G., Beckmann, C.F., Auerbach, E.J., Douaud, G., Sexton, C.E., Zsoldos, E., Ebmeier, K.P., Filippini, N., Mackay, C.E., 2014. ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *Neuroimage* 95, 232-247.
- Heinzle, J., Kahnt, T., Haynes, J.D., 2011. Topographically specific functional connectivity between visual field maps in the human brain. *Neuroimage* 56, 1426-1436.
- Hodge, M.R., Horton, W., Brown, T., Herrick, R., Olsen, T., Hileman, M.E., McKay, M., Archie, K.A., Cler, E., Harms, M.P., 2015. ConnectomeDB—sharing human brain connectivity data. *Neuroimage*.
- Jbabdi, S., Behrens, T.E., 2013. Long - range connectomics. *Annals of the New York Academy of Sciences* 1305, 83-93.
- Marcus, D.S., Harms, M.P., Snyder, A.Z., Jenkinson, M., Wilson, J.A., Glasser, M.F., Barch, D.M., Archie, K.A., Burgess, G.C., Ramaratnam, M., Hodge, M., Horton, W., Herrick, R., Olsen, T., McKay, M., House, M., Hileman, M., Reid, E., Harwell, J., Coalson, T., Schindler, J., Elam, J.S., Curtiss, S.W., Van Essen, D.C., Consortium, W.U.-M.H., 2013. Human Connectome Project informatics: quality control, database services, and data visualization. *Neuroimage* 80, 202-219.
- Meier, J.D., Aflalo, T.N., Kastner, S., Graziano, M.S., 2008. Complex organization of human primary motor cortex: a high-resolution fMRI study. *J Neurophysiol* 100, 1800-1812.

- Ongur, D., Ferry, A.T., Price, J.L., 2003. Architectonic subdivision of the human orbital and medial prefrontal cortex. *J Comp Neurol* 460, 425-449.
- Reveley, C., Seth, A.K., Pierpaoli, C., Silva, A.C., Yu, D., Saunders, R.C., Leopold, D.A., Frank, Q.Y., 2015. Superficial white matter fiber systems impede detection of long-range cortical connections in diffusion MR tractography. *Proceedings of the National Academy of Sciences* 112, E2820-E2828.
- Rilling, J.K., Glasser, M.F., Preuss, T.M., Ma, X., Zhao, T., Hu, X., Behrens, T.E., 2008. The evolution of the arcuate fasciculus revealed with comparative DTI. *Nat Neurosci* 11, 426-428.
- Robinson, E.C., Jbabdi, S., Glasser, M.F., Andersson, J., Burgess, G.C., Harms, M.P., Smith, S.M., Van Essen, D.C., Jenkinson, M., 2014. MSM: a new flexible framework for Multimodal Surface Matching. *Neuroimage* 100, 414-426.
- Salimi-Khorshidi, G., Douaud, G., Beckmann, C.F., Glasser, M.F., Griffanti, L., Smith, S.M., 2014. Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage* 90, 449-468.
- Shen, X., Tokoglu, F., Papademetris, X., Constable, R.T., 2013. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *Neuroimage* 82, 403-415.
- Van Essen, D.C., 2002. Surface - Based Atlases of Cerebellar Cortex in the Human, Macaque, and Mouse. *Annals of the New York Academy of Sciences* 978, 468-479.
- Van Essen, D.C., Dierker, D.L., 2007. Surface-based and probabilistic atlases of primate cerebral cortex. *Neuron* 56, 209-225.
- Van Essen, D.C., Jbabdi, S., Sotiropoulos, S.N., Chen, C., Dikranian, K., Coalson, T., Harwell, J., Behrens, T.E.J., Glasser, M.F., 2014. Mapping Connections in Humans and Non-Human Primates. 337-358.
- Van Essen, D.C., Smith, J., Glasser, M.F., Elam, J., Donahue, C.J., Dierker, D.L., Reid, E.K., Coalson, T., Harwell, J., 2016. The Brain Analysis Library of Spatial maps and Atlases (BALSAL) Database. *Neuroimage*.
- Wandell, B.A., Winawer, J., 2015. Computational neuroimaging and population receptive fields. *Trends in cognitive sciences*.
- Wang, J., Qiu, M., Constable, R.T., 2005. In vivo method for correcting transmit/receive nonuniformities with phased array coils. *Magn Reson Med* 53, 666-674.
- Yeo, B.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zollei, L., Polimeni, J.R., Fischl, B., Liu, H., Buckner, R.L., 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* 106, 1125-1165.
- Zilles, K., Bacha-Trams, M., Palomero-Gallagher, N., Amunts, K., Friederici, A.D., 2015. Common molecular basis of the sentence comprehension network revealed by neurotransmitter receptor fingerprints. *Cortex* 63, 79-89.