Controlling for Intra-Subject and Inter-Subject Variability in Individual-Specific Cortical Network Parcellations

Ru Kong[1], Jingwei Li[1], Nanbo Sun[1], Mert R Sabuncu[2], Alexander Schaefer[1], Xi-Nian Zuo[3,4], Avram J. Holmes[5], Simon Eickhoff[6,7], B.T. Thomas Yeo[1,8,9]

[1]ASTAR-NUS Clinical Imaging Research Centre, Department of Electrical and Computer Engineering, Singapore Institute for Neurotechnology and Memory Networks Program, National University of Singapore Singapore [2]School of Electrical and Computer Engineering, Cornell University, USA [3]CAS Key Laboratory of Behavioral Sciences and Center for Lifespan Innovation of Brain and Mind, Institute of Psychology, Beijing, China [4]University of Chinese Academy of Sciences, Beijing, China [5]Yale University, New Haven, CT, USA [6]Institute for Systems Neuroscience, Medical Faculty, Heinrich-Heine University Düsseldorf, Düsseldorf, Germany [7]Institute of Neuroscience and Medicine, Brain & Behaviour (INM-7), Research Center Jülich, Jülich, Germany [8]Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, USA [9]Centre for Cognitive Neuroscience, Duke-NUS Medical School, Singapore

Address correspondence to:

B.T. Thomas Yeo
ECE, ASTAR-NUS CIRC, SINAPSE & MNP
National University of Singapore
Email: thomas.yeo@nus.edu.sg

Abstract

Resting-state functional magnetic resonance imaging (rs-fMRI) offers the opportunity to non-invasively study brain networks within individuals. Previous individual-specific network mappings do not account for intra-subject (within-subject) variability. Therefore, intra-subject variability might be mistaken for inter-subject (between-subject) differences. Here we propose a multi-session hierarchical Bayesian model (MS-HBM) that explicitly differentiates between intra-subject and inter-subject functional connectivity variability, as well as inter-subject network spatial variability. Across three datasets, sensory-motor networks exhibited lower inter-subject, but higher intra-subject variability than association networks. Compared with other approaches, MS-HBM parcellations generalized better to new rs-fMRI and task-fMRI data. MS-HBM parcellations from a single rs-fMRI session were comparable to a recent template-matching algorithm using five sessions. Individual-specific MS-HBM cortical parcellations were highly reproducible, while capturing inter-subject network differences. Inter-subject differences in the spatial arrangement of cortical networks could predict individuals' cognition, personality and emotion, thus highlighting the importance of examining individual-specific network spatial configuration in rs-fMRI studies.

Introduction

The human cerebral cortex consists of specialized areas whose complex interactions form large-scale, spatially distributed functional networks. Recent advances in non-invasive brain imaging technologies, especially fMRI (Kwong et al., 1992; Ogawa et al., 1992), provide the opportunity to map these brain networks *in-vivo*. One prominent tool for identifying large-scale brain networks is resting-state functional connectivity (RSFC), which reflects the synchrony of rs-fMRI signals between brain regions while a subject is lying at rest without any goal-directed task (Biswal et al., 1995; Greicius et al. 2003; Fox and Raichle, 2007; Buckner et al., 2013).

Brain networks defined using RSFC have been shown to correspond well to task-evoked activation patterns (Seeley et al. 2007; Smith et al., 2009; Buckner et al., 2011; Cole et al., 2014; Krienen et al., 2014; Bertolero et al., 2015; Yeo et al., 2015a; Tavor et al., 2016). RSFC is also heritable (Glahn et al. 2010; Yang et al. 2016; Ge et al., 2017), correlates with gene expression across the cortical mantle (Hawrylycz et al. 2015; Richiardi et al. 2015; Krienen et al. 2016), and predicts individual differences in behavior (Hampson et al., 2006; van den Huevel et al., 2009; Finn et al., 2015; Rosenberg et al., 2015; Smith et al., 2015; Yeo et al., 2015b). Consequently, RSFC has been widely utilized to estimate population-level functional brain networks by averaging data across multiple subjects (Beckmann et al. 2005; Damoiseaux et al. 2006; Fox et al. 2006; Vincent et al. 2006; Dosenbach et al. 2007; Margulies et al. 2007; Calhoun et al. 2008; van den Heuvel et al. 2008; Power et al. 2011; Lee et al. 2012).

The population-level atlases of large-scale networks have provided important insights into the broad functional organization of the human brain. However, the fact that RSFC can be used to predict the behavior of individual subjects suggest the presence of behaviorally relevant inter-subject functional connectivity variability (Mueller et al., 2013; Wang et al., 2015; Laumann et al., 2015). Furthermore, the shape, location and topology of functional brain networks have been shown to vary substantially across individuals (Harrison et al., 2015; Wang et al., 2015; Langs et al., 2016; Gordon et al., 2017a; Gordon et al., 2017b). Therefore, the estimation of individual-specific brain networks could provide an important step towards precision medicine (Calhoun et al., 2009; Beckmann et al., 2009; Bellec et al., 2010; Zuo et al., 2010; Varoquaux et al., 2011; Calhoun and Adali 2012; Hacker et al., 2013; Wig et al., 2013; Wang et al., 2015; Chong et al., 2017; Gordon et al., 2017b; Braga and Buckner, 2017).

Previous approaches on estimating individual-specific brain networks either utilized a single session of rs-fMRI data or averaged data across multiple rs-fMRI sessions. In doing so, these approaches only accounted for inter-subject variability, but not intra-subject variability. However, studies have shown that the spatial patterns of inter-subject and intra-subject RSFC variability can be quite different (Mueller et al., 2013; Chen et al., 2015; Laumann et al., 2015). For example, the visual cortex exhibits relatively high intra-subject functional connectivity variability, but relatively low inter-subject functional connectivity variability (Laumann et al., 2015). Therefore, observed RSFC variability in the visual cortex might be incorrectly attributed to inter-subject spatial variability of brain networks, rather than just intra-subject sampling variability. In other words, failing to simultaneously consider both inter-subject and intra-subject variability might lead to sub-optimal network estimation, obscuring important individual-specific features of brain organization.

In this work, we proposed a multi-session hierarchical Bayesian model (MS-HBM) for deriving functional parcellations of the cerebral cortex within individual subjects. The multiple layers of the hierarchical model allowed the explicit separation of inter-subject and intra-subject functional connectivity variability. Critically, the model also provided for the differentiation of inter-subject network spatial variability and inter-subject functional connectivity variability. By applying this approach to three multi-session rs-fMRI datasets with different acquisition and preprocessing strategies, we established that the model provides robust estimates of inter-subject and intra-subject functional variability across cortex. Even though the model was defined for multi-session rs-fMRI data, an effective workaround was proposed for single-session fMRI data. The MS-HBM was also compared with three other approaches, using a held-out set of resting-state and task fMRI data collected from the same individuals for evaluation. Finally, given that individual-specific functional networks exhibit unique topological features not observed in group-level networks (Harrison et al., 2015; Laumann et al., 2015; Glasser et al., 2016; Braga & Buckner, 2017; Gordon et al., 2017a; 2017b; 2017c), we further investigated whether the spatial configuration of individual-specific cortical parcellations was behaviorally meaningful.

Methods

Overview

We proposed a multi-session hierarchical Bayesian model (MS-HBM) to estimate functional network parcellations of the cerebral cortex in individual subjects. The model distinguished between inter-subject spatial variability, as well as inter-subject and intra-subject functional connectivity variability. To assess the robustness and generalizability of our approach, the model was applied to three multi-session resting-state fMRI datasets from the Genomic Superstruct Project (GSP), Hangzhou Normal University of the Consortium for Reliability and Reproducibility (CoRR-HNU), and the Human Connectome Project (HCP), providing insights into inter-subject and intra-subject network variability. Finally, the MS-HBM was compared with three other approaches.

Multi-session fMRI datasets

The GSP test-retest dataset (Holmes et al., 2015) consisted of structural MRI and resting-state fMRI from 69 healthy young adults (ages 18 to 35). All imaging data were collected on matched 3T Tim Trio scanners (Siemens Healthcare, Erlangen, Germany) at Harvard University and Massachusetts General Hospital using the vendor-supplied 12-channel phased-array head coil. Each participant has two sessions, acquired on two different days separated by less than 6 months. One or two rs-fMRI runs were acquired per session. Each BOLD run was acquired in 3mm isotropic resolution with a TR of 3.0 seconds and lasted for 6 minutes and 12 seconds. The structural data consisted of one 1.2mm isotropic scan for each session. Details of the data collection can be found elsewhere (Holmes et al., 2015).

The CoRR-HNU multi-session dataset (Zuo et al., 2014; Chen et al., 2015) consisted of structural MRI and resting-state fMRI from 30 young healthy adults (ages 20 to 30). All imaging data were collected on 3T GE Discovery MR750 using an 8-channel head coil. Each participant was scanned a total of 10 sessions across one month (one session every three days). One rs-fMRI run was collected in each session. Each fMRI run was acquired in 3.4mm isotropic resolution with a TR of 2.0 seconds and lasted for 10 minutes. The structural data consisted of one 1mm isotropic scan for each session. Details of the data collection can be found elsewhere (Zuo et al., 2014; Chen et al., 2015).

The HCP S900 release (Van Essen et al., 2012b; Smith et al., 2013) consisted of structural MRI, resting-state fMRI and task fMRI of 881 subjects. All imaging data were collected on a custom-made Siemens 3T Skyra scanner using a multiband sequence. Each

participant has two fMRI sessions on two consecutive days. Two rs-fMRI runs were collected in each session. Each fMRI run was acquired in 2mm isotropic resolution with a TR of 0.72 seconds and lasted for 14 minutes and 33 seconds. The structural data consisted of one 0.7mm isotropic scan for each subject. Details of the data collection can be found elsewhere (Van Essen et al., 2012b; Smith et al., 2013).

Processing of GSP and CoRR-HNU data

Structural data were processed using FreeSurfer. FreeSurfer constitutes a suite of automated algorithms for reconstructing accurate surface mesh representations of the cortex from individual subjects' T1 images (Dale et al., 1999; Fischl et al., 2001; Ségonne et al., 2007). The cortical surface meshes were then registered to a common spherical coordinate system (Fischl et al. 1999a; 1999b). The GSP subjects were processed using FreeSurfer 4.5.0 (Holmes et al., 2015), while the CoRR-HNU subjects were processed using FreeSurfer 5.3.0.

Resting-state fMRI data of GSP and CoRR-HNU were initially pre-processed with the following steps: (i) removal of first 4 frames, (ii) slice time correction with the FSL package (Jenkinson et al., 2002; Smith et al., 2004), (iii) motion correction using rigid body translation and rotation with the FSL package. The structural and functional images were aligned using boundary-based registration (Greve and Fischl 2009) using the FsFast software package (http://surfer.nmr.mgh.harvard.edu/fswiki/FsFast).

Framewise displacement (FDrms) and voxel-wise differentiated signal variance (DVARS) were computed using fsl_motion_outliers. Volumes with FDrms > 0.2mm or DVARS > 50 were marked as outliers. Uncensored segments of data lasting fewer than 5 contiguous volumes were also flagged as outliers (Gordon et al., 2016). BOLD runs with more than half of the volumes flagged as outliers were removed completely. For the CoRR-HNU dataset, no session (and therefore no subject) was removed. For the GSP subjects, only one run was removed (out of a total of 222 runs). No individuals in the GSP dataset lost an entire session, and therefore, all subjects were retained.

Linear regression using multiple nuisance regressors was applied. Nuisance regressors consisted of global signal, six motion correction parameters, averaged ventricular signal, averaged white matter signal, as well as their temporal derivatives (18 regressors in total). The flagged outlier volumes were ignored during the regression procedure. The data were interpolated across censored frames using least squares spectral estimation of the values at censored frames (Power et al., 2014). Finally, a band-pass filter ($0.009\ \text{Hz} \leq f \leq 0.08\ \text{Hz}$) was applied.

The preprocessed fMRI data was projected onto the FreeSurfer fsaverage6 surface space (2mm vertex spacing). The projected fMRI data was smoothed using a 6mm full-width half-maximum kernel and then downsampled onto fsaverage5 surface space (4mm vertex spacing). Smoothing on the fsaverage6 surface, rather than in the volume minimized the blurring of fMRI signal across sulci.

Processing of HCP data

Details of the HCP data collection and preprocessing can be found elsewhere (HCP S900 manual; Van Essen et al. 2012b; Glasser et al. 2013; Smith et al. 2013). Of particular importance is that the rs-fMRI data has been projected to the fsLR surface space (Van Essen et al. 2012a), smoothed by 2mm and denoised with ICA-FIX (Salimi-Khorshidi et al. 2014; Griffanti et al., 2014).

However, recent studies have shown that ICA-FIX does not fully eliminate global and head-motion related artifacts (Burgess et al., 2016; Siegel et al., 2016). Therefore, further processing steps were performed on the rs-fMRI data in fsLR surface after ICA-FIX denoising, which included nuisance regression, motion censoring and interpolation, and band-pass filtering. Volumes with FDrms > 0.2mm or DVARS > 75, as well as uncensored segments of data lasting fewer than 5 contiguous volumes were flagged as outliers. BOLD runs with more than half the volumes flagged as outliers were completely removed. Consequently, 56 subjects were removed. Furthermore, for this work, only subjects with all four runs remaining (N = 676) were considered.

Nuisance regression utilized regressors consisting of global signal, six motion parameters, averaged ventricular signal, averaged white matter signal, and their temporal derivatives (18 regressors in total). The outlier volumes were ignored during the regression procedure. The data were interpolated across censored frames using least squares spectral estimation (Power et al., 2014). A band-pass filter ($0.009\ \text{Hz} \leq f \leq 0.08\ \text{Hz}$) was then applied to the data. Finally, spatial smoothing was applied by iteratively averaging the data at each surface mesh vertex with its neighbors four times.

Population-level parcellation and functional connectivity profiles

We have previously developed an approach to derive a population-level parcellation of the cerebral cortex into large-scale resting-state networks (Yeo et al., 2011). The cortical networks were defined as sets of cortical regions with similar corticocortical functional connectivity profiles. Here we applied the same approach to the GSP, CoRR-HNU and HCP

datasets. Our previous analyses (Yeo et al., 2011) identified 7 and 17 networks to be particularly stable. For simplicity, we will only consider 17 networks. Details of this approach have been previously described (Yeo et al., 2011). For completeness, we briefly described its application to the current datasets.

Recall that the preprocessed fMRI data from the CoRR-HNU and GSP subjects have been projected onto the fsaverage5 surface meshes. The fsaverage5 surface meshes consisted of 18715 cortical vertices. We defined the connectivity profile of a cortical region (vertex) to be its functional coupling to 1175 regions of interest (ROIs). The 1175 ROIs consisted of single vertices uniformly distributed (16mm spacing) across the fsaverage5 surface meshes. For each rs-fMRI run of each subject, the Pearson's product moment correlation between the fMRI time series at each spatial location (18715 vertices) and the 1175 ROIs were computed. The 18715 x 1175 correlation matrix were then binarized by keeping the top 10% of the correlations to obtain the final functional connectivity profiles. The outlier volumes (flagged during preprocessing) were ignored when computing the correlations.

In the case of the HCP dataset, the preprocessed fMRI data have been projected onto the fsLR surface space. The fsLR_32K surface meshes consisted of 59412 cortical vertices. We defined the connectivity profile of a cortical region (vertex) to be its functional coupling to 1483 ROIs. The 1483 ROIs consisted of single vertices uniformly distributed (14 mm apart) across the fsLR_32K surface meshes. For each rs-fMRI run of each subject, the Pearson's product moment correlation between the fMRI time series at each spatial location (59412 vertices) and the 1483 ROIs were computed. The 59412 x 1483 correlation matrix were then binarized by keeping the top 10% of the correlations to obtain the final functional connectivity profile. The outlier volumes (flagged during preprocessing) were ignored when computing the correlations.

To obtain a population-level parcellation from a group of subjects, each vertex's connectivity profiles were averaged across all BOLD runs of all subjects. The averaged connectivity profiles were clustered using a mixture of von Mises–Fisher distributions (Lashkari et al., 2010; Yeo et al., 2011). The algorithm operated by first randomly assigning the vertices (18715 in the GSP and CoRR-HNU datasets, or 59412 in the HCP dataset) to different networks. The algorithm then iterated between two steps (E-step and M-step) until convergence. In the M-step, the algorithm computed a network-level connectivity profile based on vertices assigned to the same network. In the E-step, the algorithm re-assigned the network membership of vertices based on the similarity between each vertex's connectivity profile and the network-level connectivity profile. The clustering algorithm was repeated

1000 times with different random initializations and the estimate with the best model likelihood was selected.

Multi-session hierarchical Bayesian model (MS-HBM)

The previous section described an approach to estimate a population-level parcellation from a group of subjects. Figure 1 illustrates the MS-HBM model for estimating individual-specific cerebral cortex parcellations using multi-session fMRI data. Some of the model parameters (e.g., inter-subject variability) must be estimated from a training set of subjects. A new subject (possibly from another dataset) could then be parcellated without access to the original training data. Even though the model was defined on multi-session fMRI data, an effective workaround was provided for single-session fMRI data. The exact mathematical model is found in Supplemental Methods S1. Here we provide the intuition behind this model.

Let $X_n^{s,t}$ denote the (binarized) functional connectivity profile of cortical vertex n from session t of subject s. For example, Figure 1 illustrates the binarized functional connectivity profile for a posterior cingulate cortex vertex ($X_{PCC}^{1,1}$) and a precuneus vertex ($X_{pCun}^{1,1}$) from the first session of the first subject. Based on the connectivity profiles of all vertices from all sessions of a single subject, the goal was to assign a network label l_n^s for each vertex n of subject s. Even though a vertex's connectivity profiles were unlikely to be the same across different fMRI sessions, the vertex's network label was assumed to be the same across sessions.

Consistent with previous work (Yeo et al., 2011), the von Mises–Fisher mixture model was utilized to encourage brain locations with similar functional connectivity profiles to be assigned the same network label (illustrated by arrow from network label l_n^s to connectivity profile $X_n^{s,t}$ in Figure 1). For example, the connectivity profiles of PCC ($X_{PCC}^{s,t}$) and precuneus ($X_{pCun}^{s,t}$) were very similar, so they were more likely to be grouped into the same network (i.e., default mode network or DMN).

However, unlike the group averaged connectivity profiles, the functional connectivity profiles of individual subjects are generally very noisy. If the connectivity profiles of PCC and pCun were too noisy, the mixture model might not assign them to the same network. Therefore, an additional spatial smoothness prior was incorporated. More specifically, the spatial smoothness prior V (Potts model) encouraged neighboring vertices (e.g., PCC and pCun) to be assigned to the same network.

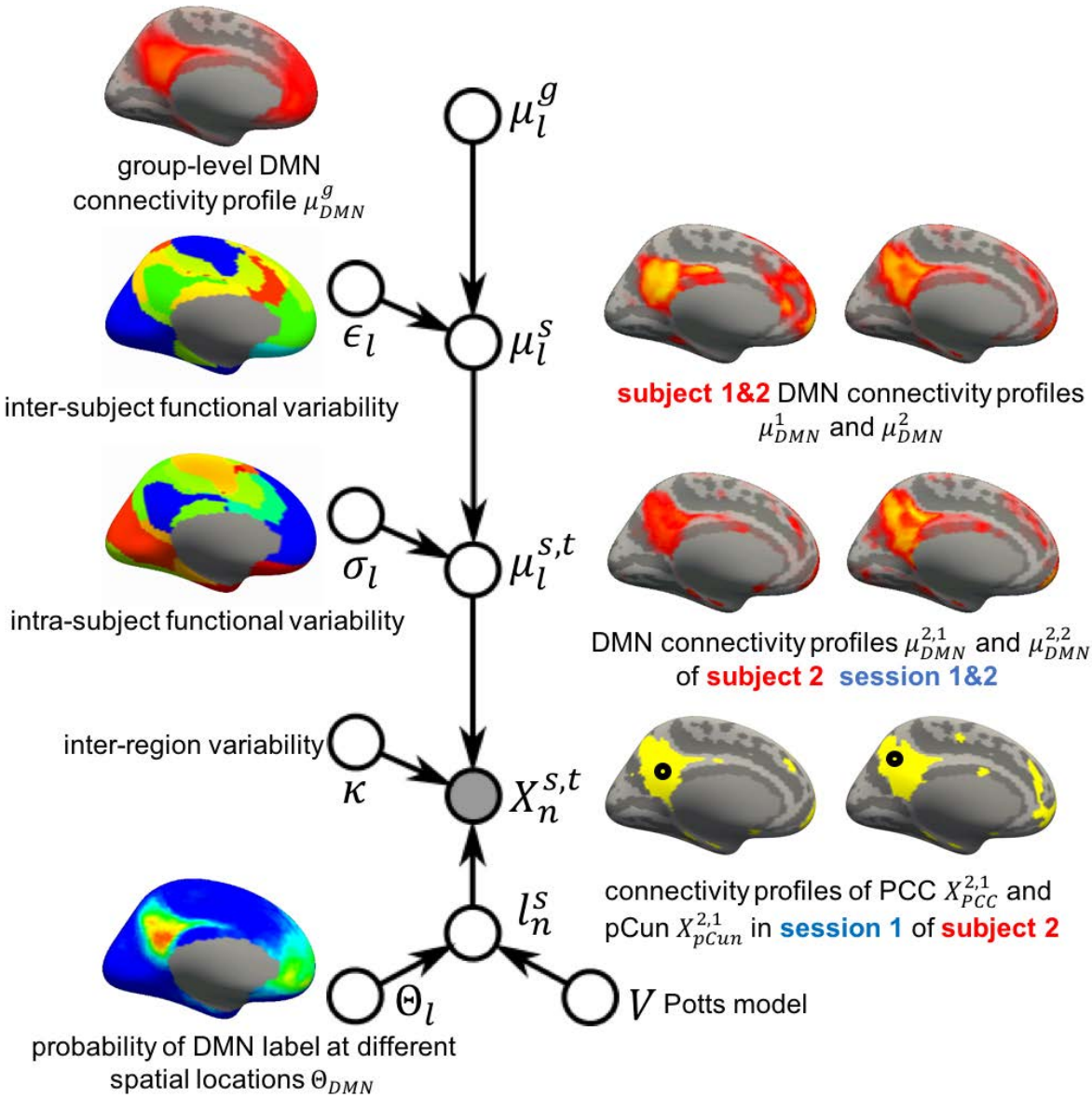

Figure 1. Multi-session hierarchical Bayesian model (MS-HBM) of individual-specific cortical parcellation. Given RSFC profile $X_n^{s,t}$ at brain location n for subject s during rs-fMRI session t, the goal is to estimate the network label l_n^s. μ_l^g is the group-level RSFC profile of network l, μ_l^s is the subject-level RSFC profile of network l, and $\mu_l^{s,t}$ is the subject-level RSFC profile of network l during session t. Inter-subject and intra-subject RSFC profile variabilities are captured by ϵ_l and σ_l. κ captures inter-region RSFC variability, e.g., posterior cingulate cortex (PCC) and precuenus (pCun) might belong to the same network, but might exhibits slightly different RSFC profiles. Finally, Θ_l captures inter-subject variability in the spatial distribution of networks, while smoothness prior V encourages network labels to spatially smooth. See text for details.

To model inter-subject spatial variability, the spatial prior $\Theta_{l,n}$ denote the probability of network l occurring at a particular spatial location n. As an example, the spatial variability map of the default network (Θ_{DMN}) is shown in Figure 1 (bottom left), where warm color

indicates high probability and cool color indicates low probability. Both PCC and pCun had high prior probabilities of being assigned to the default network.

To model inter-subject functional connectivity variability, let μ_l^g denote the group-level functional connectivity profile of network l. For example, Figure 1 (top left) illustrates the group-level DMN connectivity profile (μ_{DMN}^g). Let μ_l^s denote the functional connectivity profile of network l and subject s. For example, Figure 1 (top right) illustrates the DMN connectivity profiles of two different subjects (μ_{DMN}^1 and μ_{DMN}^2). The parameter ϵ_l controlled how much the individual-specific network connectivity profile μ_l^s can deviate from the group-level network connectivity profile μ_l^g, and therefore represented the amount of inter-subject functional connectivity variability. For example, Figure 1 illustrates the inter-subject connectivity variability ϵ for the 17 networks considered in this paper. Hotter colors indicate higher connectivity variability. The default network was colored green, which indicated an intermediate amount of inter-subject functional connectivity variability.

To model intra-subject functional connectivity variability, let $\mu_l^{s,t}$ denote the functional connectivity profile of network l and subject s during session t. For example, Figure 1 illustrates the default network connectivity profiles of subject 2 during sessions 1 and 2 ($\mu_{DMN}^{2,1}$ and $\mu_{DMN}^{2,2}$). The parameter σ_l controlled how much the session-specific network connectivity profile $\mu_l^{s,t}$ could deviate from the individual-specific network connectivity profile μ_l^s, and therefore represented the amount of intra-subject variability. For example, the intra-subject functional connectivity variability σ for the 17 networks are shown in Figure 1. Hotter colors indicate higher connectivity variability. The default network was colored blue, which indicated low intra-subject functional connectivity variability.

The functional connectivity profile $\mu_l^{s,t}$ could be thought of as the representative connectivity profile of vertices belonging to network l of subject s during session t. However, the connectivity profiles of two regions belonging to the same network (e.g., $X_{pCun}^{s,t}$ and $X_{PCC}^{s,t}$) might exhibit slightly different connectivity profiles. Suppose vertex n is assigned to network l. The parameter κ controlled how much the connectivity profile $X_n^{s,t}$ of vertex n from session t of subject s could deviate from the individual-specific session-specific network connectivity profile $\mu_l^{s,t}$. For simplicity, κ was assumed to be the same across networks and all subjects.

The group-level network connectivity profiles μ_l^g, the inter-subject functional connectivity variability ϵ_l, the intra-subject functional connectivity variability σ_l, the spatial

smoothness prior V and the inter-subject spatial variability prior Θ_l could be estimated from a dataset of multiple subjects with multi-session fMRI data. Conditioned on the estimated model parameters ($\mu_l^g, \epsilon_l, \sigma_l, V, \theta_l$), the parcellation of a new subject could then be estimated. Here we utilized a variational Bayes Expectation-Maximization (VBEM) algorithm to learn the model parameters from the training data and to estimate individual-specific parcellations. Details of the VBEM algorithm can be found in Supplementary Methods S2.

Experimental setup

For the purpose of subsequent experiments, the GSP dataset was divided into training (N = 37) and validation (N = 32) sets. The CoRR-HNU dataset (N = 30) was kept unchanged. The HCP dataset was divided into training (N = 40), validation (N = 40) and test (N = 596) sets.

For simplicity, different fMRI runs within the same session were treated as data from different sessions. For example, each HCP subject underwent two fMRI sessions on two consecutive days. Within each session, there were two rs-fMRI runs. For the purpose of our analyses, we treated each HCP subject as having four sessions of data. Future work might differentiate between intra-session and inter-session variability.

Although the MS-HBM was formulated for multi-session fMRI data, most studies only collect a single run of fMRI data. We considered the ad-hoc approach of splitting the single fMRI run into two and treating the resulting runs as two separate sessions. Our evaluations (see Results section) suggest that this workaround worked surprisingly well.

Characterizing inter-subject and intra-subject network variability

The group-level parcellation algorithm was applied to the GSP training dataset. The resulting group-level parcellation was then used to initialize the estimation of the group-level network connectivity profiles μ_l^g, the inter-subject functional connectivity variability ϵ_l, the intra-subject functional connectivity variability σ_l, and the inter-subject spatial variability prior Θ_l. For this analysis, the spatial smoothness prior V was ignored. The estimated inter-subject and intra-subject functional connectivity variability maps, as well as the inter-subject spatial variability maps were visualized in Figures 2 and 3. The procedure was repeated for the CoRR-HNU dataset and HCP training set, allowing us to evaluate whether inter-subject and intra-subject variability could be reliably estimated across datasets.

Given that inter-subject and intra-subject variability might differ across sites, an important consideration was whether model parameters estimated from one dataset could be generalized to another dataset. Therefore, the MS-HBM was also trained on the GSP dataset and applied to the CoRR-HNU dataset. The training procedure was the same as before, except that the GSP validation set was also used to tune the spatial smoothness prior V. The individual-specific parcellations were independently inferred using sessions 1-5 and sessions 6-10 of the CoRR-HNU dataset. Therefore, there were two individual-specific parcellations for each subject based on data from two independent sets of five sessions.

To evaluate the reproducibility of individual-specific parcellations, the Dice coefficient was computed for each network from the two parcellations of each subject. The Dice coefficients were then averaged across all networks and all subjects to provide an overall measure of intra-subject parcellation reproducibility. To evaluate inter-subject parcellation similarity, for each pair of subjects, the Dice coefficient was computed for each network. Since there were two parcellations for each subject, there were a total of four Dice coefficients for each network, which were then averaged. The Dice coefficients were then averaged across all networks and all pairs of subjects to provide an overall measure of inter-subject parcellation similarity.

Quantitative comparison with alternative approaches

We compared the MS-HBM approach with three alternative approaches. The first approach was to apply the population-level parcellation (Yeo et al., 2011) to individual subjects. We will refer to this approach as Yeo2011. Recall that the population-level parcellation algorithm iteratively computed a network connectivity profile based on vertices assigned to the same network (M-step) and then re-assigned the network membership of vertices based on the similarity between each vertex's connectivity profile and the network connectivity profile (E-step). Using the network connectivity profiles from the population-level parcellation, we can estimate networks in an individual subject by assigning a network label to each vertex based on the similarity between the vertex's connectivity profile (for that subject) and the population-level network connectivity profile (i.e., E-step). Since this approach is similar to the ICA back-projection algorithm (Calhoun et al., 2009; Beckmann et al., 2009; Filippini et al., 2009; Zuo et al., 2010; Calhoun and Adali 2012), we will refer to this algorithm, which is our second alternative, as YeoBackProject. Finally, we also implemented the influential individual-parcellation algorithm of Gordon and colleagues (Gordon et al., 2017a; Gordon et al., 2017b), where the binarized functional connectivity map

of each cortical vertex was matched to binarized network templates derived from the group-level parcellation. We refer to this approach as Gordon2017.

The model parameters of all algorithms were estimated from the GSP dataset and then utilized to infer the parcellation of each CoRR-HNU subject using one or more fMRI sessions. Varying the number of sessions allowed us to evaluate the amount of data necessary for the algorithms to work well. Because the HCP data were in a different surface space from the GSP data, we could not apply the GSP model parameters to the HCP subjects. Instead, the model parameters of all algorithms were re-estimated from the HCP training and validation sets, and then utilized to infer the parcellation of each subject in the HCP test set using the first resting-state fMRI run from the first session.

If an individual-specific parcellation captured the system-level organization of the individual's cerebral cortex, then each network should have homogeneous function and connectivity. Therefore, the following resting-state connectional homogeneity and task functional inhomogeneity measures were used as parcellation evaluation metrics (Gordon et al., 2015; Gordon et al., 2017c; Schaefer et al., in press):

1. Resting-state connectional homogeneity. Resting-state connectional homogeneity was computed by averaging the Pearson's correlations between the resting-state fMRI time courses of all pairs of vertices within each network. The average correlations are then averaged across all networks while accounting for network size:

$$\frac{\sum_{l=1}^{L} \rho_l |l|}{\sum_{l=1}^{L} |l|}, \qquad (1)$$

   where ρ_l is the resting-state homogeneity of network l and $|l|$ is the number of vertices within network l (Schaefer et al., in press). For each subject from CoRR-HNU (N = 30) and HCP test set (N = 596), we used one or more sessions to infer the individual-specific parcellation and computed the resting-state homogeneity of the individual-specific parcellation with the remaining sessions.

2. Task functional inhomogeneity. The HCP task-fMRI data consisted of seven functional domains: social cognition, motor, gambling, working memory, language processing, emotional processing and relational processing, each with multiple task contrasts (Barch et al., 2013). For a given task contrast, task inhomogeneity was defined as the standard deviation of (activation) z-values within each network. A lower standard deviation

indicates higher functional homogeneity within the network. The standard deviations are averaged across all networks while accounting for network size:

$$\frac{\sum_{l=1}^{L} std_l |l|}{\sum_{l=1}^{L} |l|}, \qquad (2)$$

where std_l is the standard deviation of task activation z-values for network l and $|l|$ is the number of vertices in parcel l (Gordon et al., 2017b; Schaefer et al., in press).

For each subject in the HCP test set (N = 596), the first resting-state fMRI run from the first session was used to infer the individual-specific parcellation. The individual-specific parcellation was then utilized to evaluate task inhomogeneity for each task contrast (Eq. (2)) and then averaged across all contrasts within a functional domain, resulting in a single functional inhomogeneity measure per functional domain. The number of task contrasts per functional domain ranged from three for the emotion domain to eight for the working memory domain. When comparing between parcellations, the inhomogeneity metric (Eq. (2)) was averaged across all contrasts within a functional domain before a paired-sample t-test (dof = 595) was performed for each functional domain.

Behavioral relevance of individual-specific MS-HBM parcellations

Individual-specific MS-HBM parcellations were estimated for each HCP test subject (N = 596) using all four rs-fMRI runs. To test the behavioral relevance of individual-specific MS-HBM parcellations, we selected 58 behavioral phenotypes measuring cognition, personality and emotion (Table S1). 17 subjects were excluded from further analyses because they did not have all behavioral phenotypes, resulting in a final set of 579 subjects.

Kernel regression (Murphy et al., 2012) was utilized to predict each behavioral phenotype in individual subjects. Suppose y is the behavioral measure (e.g., fluid intelligence) and l is the individual-specific parcellation of a test subject. In addition, suppose y_i is the behavioral measure (e.g., fluid intelligence) and l_i is the individual-specific parcellation of the i-th training subject. Then kernel regression would predict the behavior of the test subject as the linear combination of the behaviors of the training subjects: $y = \sum_{i \in \text{training set}} y_i \, \text{Similarity}(l_i, l)$. Here, $\text{Similarity}(l_i, l)$ is set to be the Dice coefficient for each network, averaged across 17 networks. Therefore, kernel regression assumes that subjects with more similar parcellations have similar behavioral measures.

In practice, we included a regularization term (i.e., kernel ridge regression) estimated via an inner-loop cross-validation procedure (Murphy et al., 2012). More specifically, we performed 20-fold cross-validation for each behavioral phenotype. Care was taken so that family members were not split between folds. For each test fold, inner-loop cross validation was applied to the remaining 19 folds to determine the best regularization parameter. The optimal regularization parameter from the 19 folds was then used to predict the behavioral phenotype in the test fold. Accuracy was measured by correlating the predicted and actual behavior across all subjects within the test fold (Finn et al., 2015), resulting in 20 correlation accuracies for each behavior. To test whether the predictions were statistically better than chance, the accuracies were averaged across all behaviors and a corrected resampled t-test (dof = 19) was performed (Nadeau and Bengio, 2000; Bouckart and Frank, 2004).

Finally, we note that certain behavioral measures are correlated with motion (Siegel et al., 2016). Therefore, age, sex and motion were regressed from the behavioral data before kernel ridge regression. To prevent information from the training data to leak to the test data, the nuisance regression was performed on the training folds and the regression coefficients were applied to the test fold.

Code availability

Code for this work is freely available at the github repository maintained by the Computational Brain Imaging Group (https://github.com/ThomasYeoLab/CBIG). More specifically, the GSP and CoRR-HNU datasets were preprocessed using an in-house pipeline that can be downloaded here (GITHUB_LINK_TO_BE_ADDED). The group-level parcellation code (Yeo et al., 2011) can be downloaded here (https://github.com/ThomasYeoLab/CBIG/tree/master/stable_projects/brain_parcellation/Yeo2011_fcMRI_clustering). Finally, the individual-specific parcellation code can be downloaded here (GITHUB_LINK_TO_BE_ADDED).

Results

Sensory-motor networks exhibit lower inter-subject, but higher intra-subject, functional connectivity variability than association networks.

Figure 2A shows the 17-network population-level parcellation estimated from the GSP dataset (Yeo et al., 2011). The 17 networks were divided into eight groups (Visual, Somatomotor, Dorsal Attention, Salience/Ventral Attention, Limbic, Control, Default and TempPar), which broadly corresponded to major networks discussed in the literature (Yeo et al., 2015b). The 17 networks are referred to as "Default A", "Default B" and so on (Figure 2A).

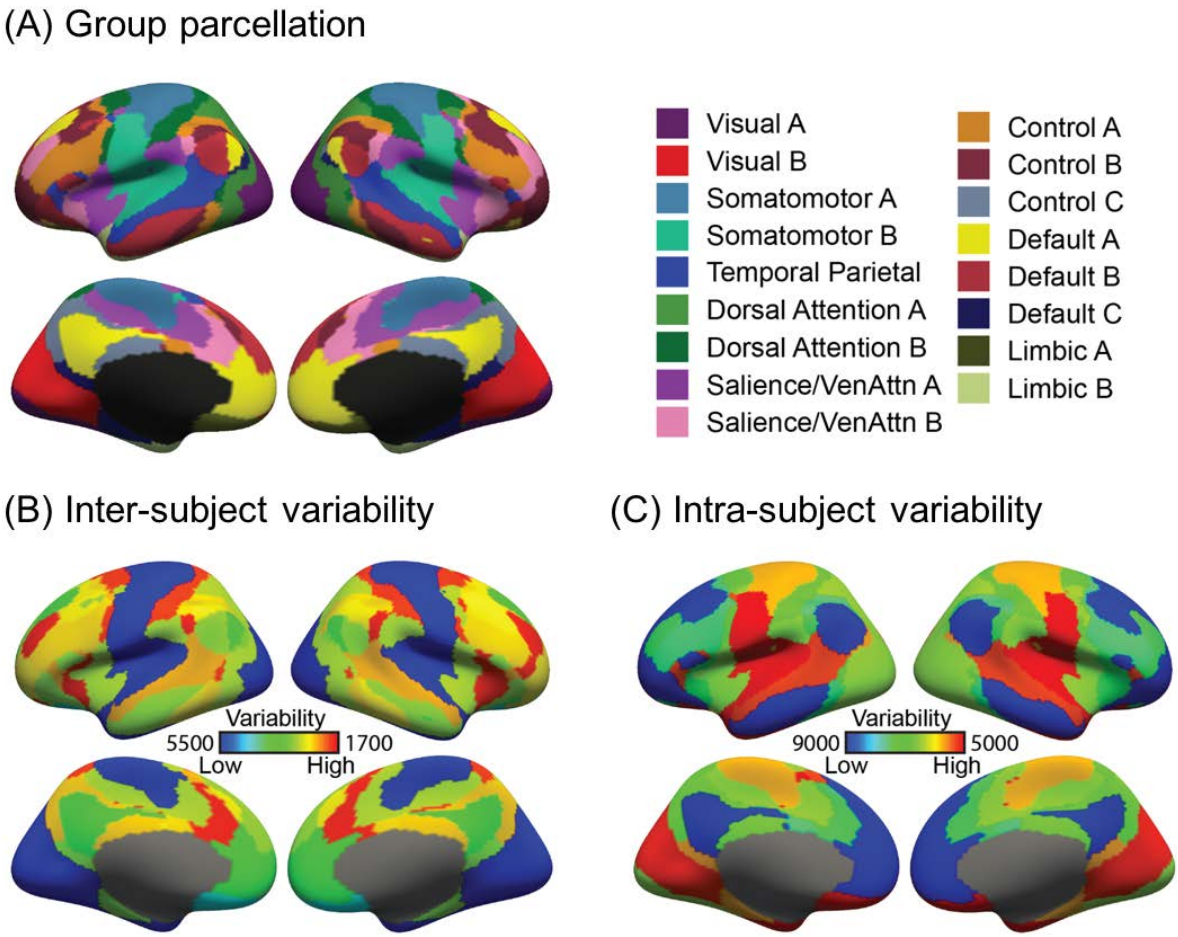

Figure 2. Sensory-motor networks exhibit lower inter-subject, but higher intra-subject, functional connectivity variability than association networks in the GSP training set. (A) 17-network group-level parcellation. (B) Inter-subject functional connectivity variability for different cortical networks. (C) Intra-subject functional connectivity variability for different cortical networks. Results were replicated in the CoRR-HNU (Figure S1) and HCP (Figure S2) datasets. Note that (B) and (C) correspond to the ϵ_l and σ_l parameters in Figure 1.

The GSP population-level parcellation was replicated in the CoRR-HNU (Figure S1) and HCP (Figure S2) datasets, although there were some interesting differences with the HCP dataset, possibly due to distinct acquisition sequences. For example, the Limbic networks (A and B) from the GSP population-level parcellation (Figure 1) were absorbed into the Default networks (A and B) in the HCP population-level parcellation (Figure S2A). Instead, there were two additional networks in the HCP population-level parcellation: Visual C and Auditory networks. The Visual C network (Figure S2A) might correspond to the foveal representation within the primary visual cortex, while the Auditory network (Figure S2A) appeared to have split off from the Somatomotor B network in the GSP population-level parcellation (Figure 2A).

Figure 2B shows the inter-subject functional connectivity variability map estimated from the GSP dataset. Sensory-motor networks exhibited lower inter-subject functional connectivity variability than association networks. More specifically, Somatomotor (A and B) and Visual (A and B) networks were the least variable, while Salience/Ventral Attention Network B was the most variable. The results were largely consistent in the CoRR-HNU (Figure S1) and HCP (Figure S2) datasets, although there were some notable differences. For example, the Somatomotor B network exhibited low variability in both the GSP and HCP datasets, but intermediate variability in the CoRR-HNU dataset.

Figure 2C shows the intra-subject functional connectivity variability map estimated from the GSP dataset. In general, association networks exhibited lower intra-subject functional connectivity variability than sensory-motor networks. More specifically, Default networks (A and B) were the least variable, while Somatomotor networks (A and B) and Visual network B were the most variable. The results were largely consistent in the CoRR-HNU (Figure S1) and HCP (Figure S2) datasets, although there were some interesting differences. Of particular note is that the Visual Network B exhibited high intra-subject functional connectivity variability in the GSP dataset, but low or intermediate functional connectivity variability in the CoRR-HNU and HCP datasets (see Discussion).

It is worth noting that the values in Figure 2C are much larger than Figure 2B, suggesting that intra-subject functional connectivity variability is much lower than inter-subject functional connectivity variability. These results are replicated in the CoRR-HNU (Figure S1) and HCP (Figure S2) datasets.

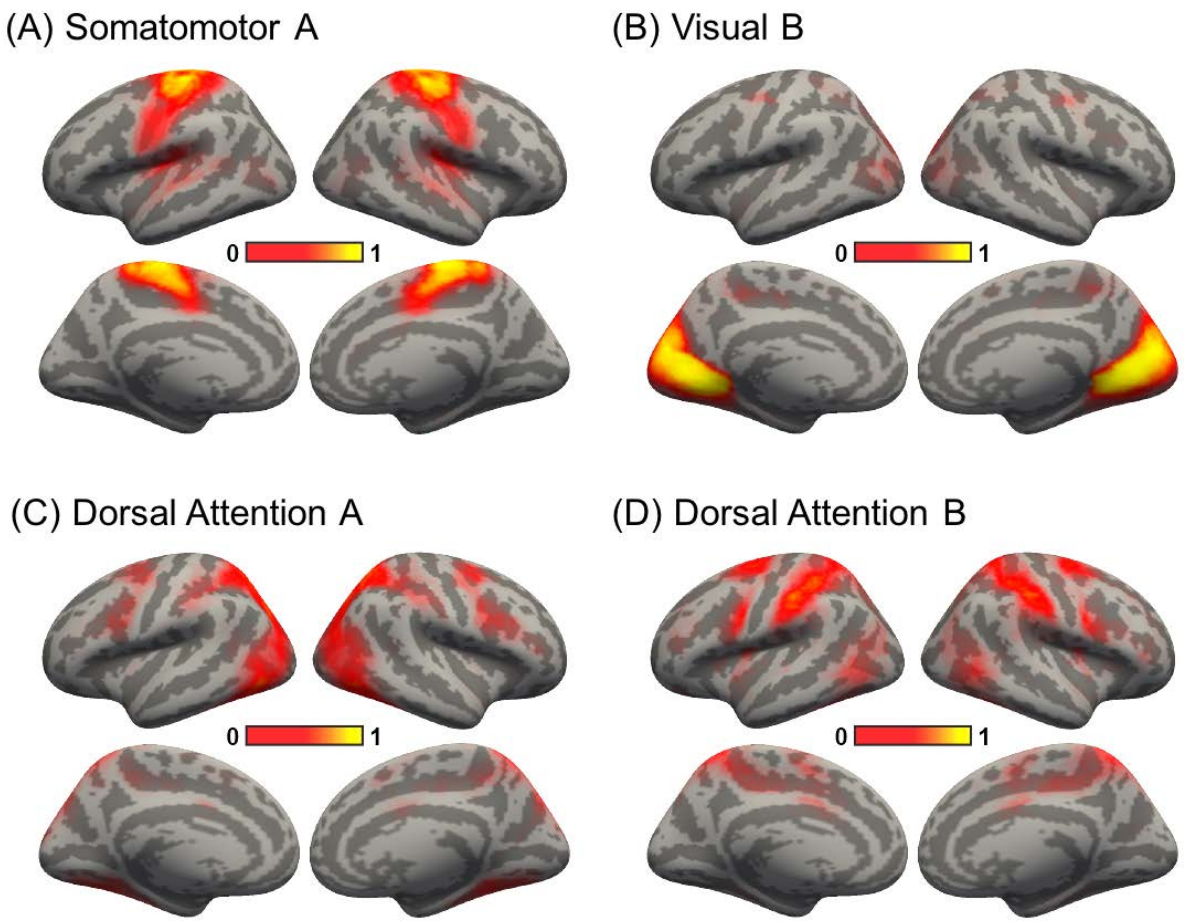

Figure 3. Sensory-motor networks are less spatially variable than association networks across subjects in the GSP training set. Spatial probability maps of (A) Somatomotor network A, (B) Visual network B, (C) Dorsal Attention network A, and (D) Dorsal Attention network B. A higher value (bright color) at a spatial location indicates high probability of a network appearing at that spatial location. Results were replicated in the CoRR-HNU (Figure S3) and HCP (Figure S4) datasets. Note that this corresponds to the Θ_l parameter in Figure 1.

Sensory-motor networks are less spatially variable than association networks across subjects

Our approach allowed the modeling of both inter-subject functional connectivity and network spatial variability. Like inter-subject functional connectivity variability, the sensory-motor networks were found to be less spatially variable than association networks across subjects. For example, Figure 3 shows the inter-subject spatial variability maps of four representative networks from the GSP dataset. Yellow color at a spatial location indicates that across subjects, there is a high probability of the network appearing at that spatial location, suggesting low inter-subject spatial variability. The Somatomotor network A and Visual network B showed higher probabilities (more yellow color) than the Dorsal Attention networks (Figure 3), suggesting that Somatomotor network A and Visual network B

exhibited lower inter-subject spatial variability than Dorsal Attention networks. These results were consistent in the CoRR-HNU and HCP datasets (Figures S3 and S4).

Individual-specific MS-HBM parcellations exhibit high intra-subject reproducibility and low inter-subject similarity

To assess intra-subject reproducibility and inter-subject similarity, our model (Figure 1) was trained on the GSP dataset and applied to the CoRR-HNU dataset. Individual-specific parcellations were generated by using sessions 1-5 and sessions 6-10 separately for each subject. Figure 4 shows the parcellations of four representative subjects. The 17 networks were present in all individual-specific parcellations, but the shapes, sizes and topologies were varied across subjects.

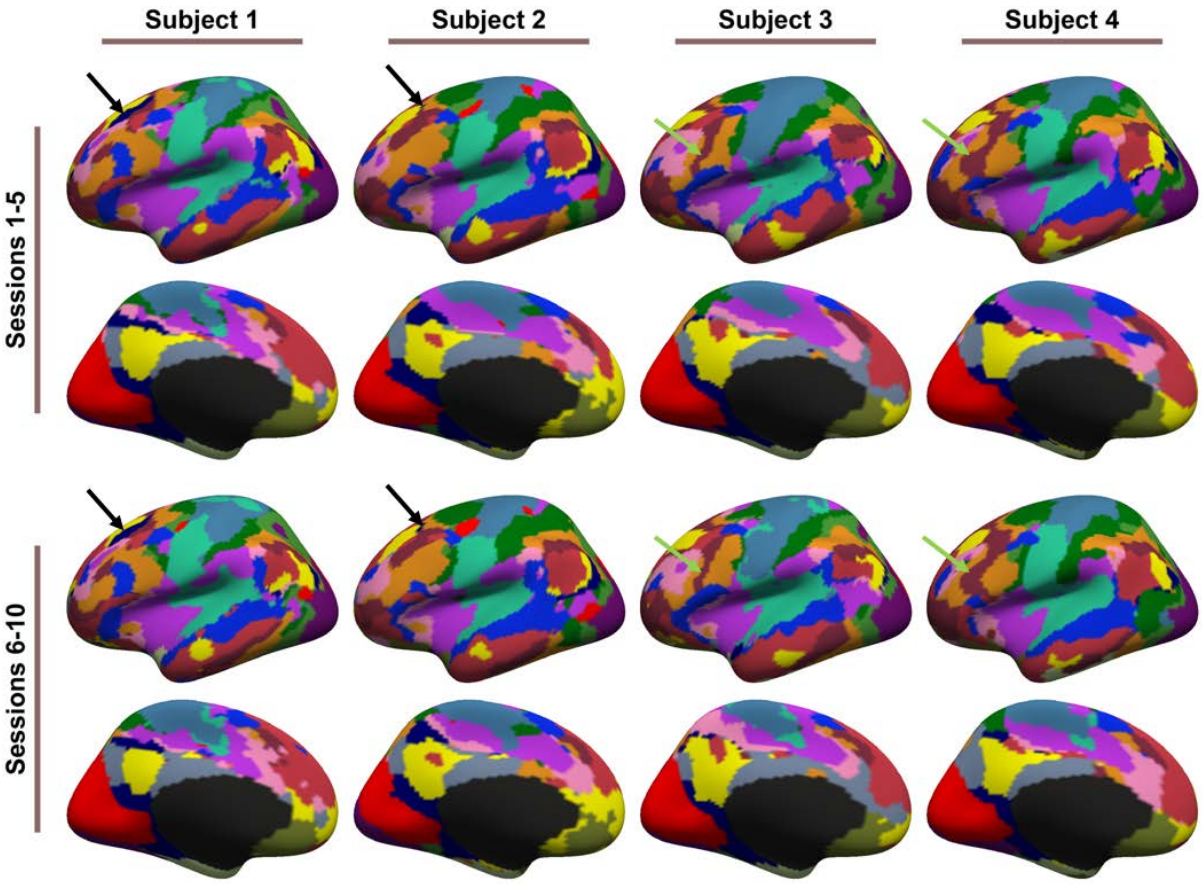

Figure 4. 17-network parcellations were estimated using sessions 1-5 and sessions 6-10 separately for each subject from the CoRR-HNU dataset. Parcellations of four representative subjects are shown here. Black and blue arrows indicate individual-specific parcellation features. Right hemisphere parcellations are shown in Figure S5.

For example, the Default C (dark blue) network exhibited a dorsal lateral prefrontal component for certain subjects (black arrows in Figure 4), but was missing in other subjects. As another example, the lateral prefrontal component of the Control A (orange) network was separated into two separate components by the Control B (brown) network (green arrows in Figure 4). These features were mostly replicated across sessions.

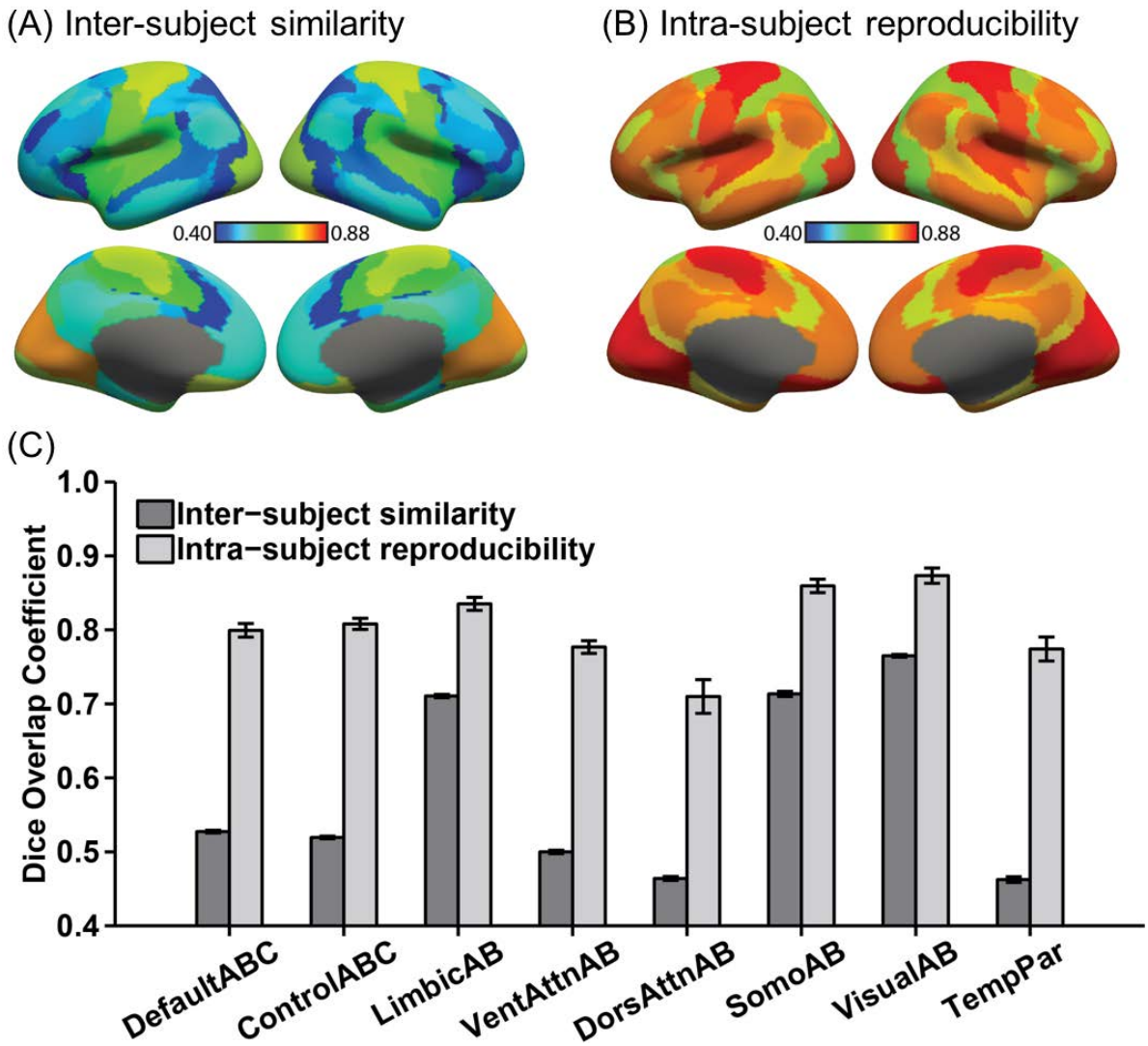

Figure 5. Individual-specific MS-HBM parcellations show high within-subject reproducibility (overlap = 81.59%) and low across-subject similarity (overlap = 59.36%) in the CoRR-HNU dataset. (A) Inter-subject spatial similarity for different networks. (B) Intra-subject reproducibility for different networks. Warm color indicates higher overlap. Cool color indicates lower overlap. (C) Quantification of inter-subject similarity and intra-subject reproducibility for different networks. "VentAttnAB" corresponds to Salience/Ventral Attention networks A and B. "SomoAB" corresponds to Somatomotor networks A and B.

Figure 5A shows the across-subject spatial similarity (Dice coefficient) of individual-specific parcellations. A higher value (hot color) indicates greater inter-subject agreement.

Figure 5B shows the within-subject reproducibility (Dice coefficient) of individual-specific parcellations. A higher value (hot color) indicates greater inter-session agreement within subjects. Further quantification is shown in Figure 5C, where the Dice coefficients were averaged across sub-networks.

Across all networks, intra-subject reproducibility was greater than inter-subject similarity. Compared with association networks, the Somatomotor networks (A and B) and Visual networks (A and B) were more spatially similar across subjects, but also exhibited greater within subject inter-session reproducibility. Among association networks, the Dorsal Attention networks (A and B) and Salience/Ventral Attention networks (A and B) were the least similar across subjects, while also exhibiting the least within subject inter-session reproducibility. Overall, our parcellation model achieved 81.6% intra-subject reproducibility and 59.4% inter-subject similarity.

Individual-specific networks generated by MS-HBM exhibit higher resting-state homogeneity than other approaches

Individual-specific parcellations were estimated using one rs-fMRI session from the CoRR-HNU dataset and HCP test set. The resting-state homogeneity of the parcellations were evaluated in the leave-out sessions (Figure 6A). Across both CoRR-HNU and HCP datasets, the group-level parcellation (Yeo2011) achieved the worst resting-state homogeneity. On the other hand, the MS-HBM performed better than the other approaches. In the CoRR-HNU dataset, compared with Yeo2011, YeoBackProject and Gordon2017, the MS-HBM achieved a homogeneity improvement of 16.6% ($p = 3.23e-21$), 5.32% ($p = 4.47e-18$) and 6.88% ($p = 1.23e-17$) respectively. In the HCP dataset, compared with Yeo2011, YeoBackProject and Gordon2017, the MS-HBM achieved an improvement of 9.8% ($p < 5e-324$), 9.54% ($p < 5e-324$) and 5.74% ($p < 5e-324$) respectively.

Individual-specific parcellations were estimated with increasing number of rs-fMRI sessions using the CoRR-HNU dataset. The resting-state homogeneity of the parcellations were evaluated in the leave-out sessions (Figure 6B). Not surprisingly, performance of the Yeo2011 group-level parcellation remained constant regardless of the amount of data. The remaining three approaches (YeoBackProject, Gordon2017 and MS-HBM) exhibited higher homogeneity with increased number of sessions. Critically, the improvement of our model over the other approaches grew with the inclusion of additional fMRI sessions. For example, as the number of sessions was increased from two to three to four to five, our approach achieved improvement of 5.44%, 5.9%, 6.13% and 6.38% respectively over Gordon2017.

Interestingly, the improvement of our approach over Gordon2017 was largest when only one rs-fMRI session was utilized (6.88%). Furthermore, using just one fMRI sessions (10 min), our algorithm was able to match the homogeneity achieved with the Gordon2017 approach that used five fMRI sessions (50 min).

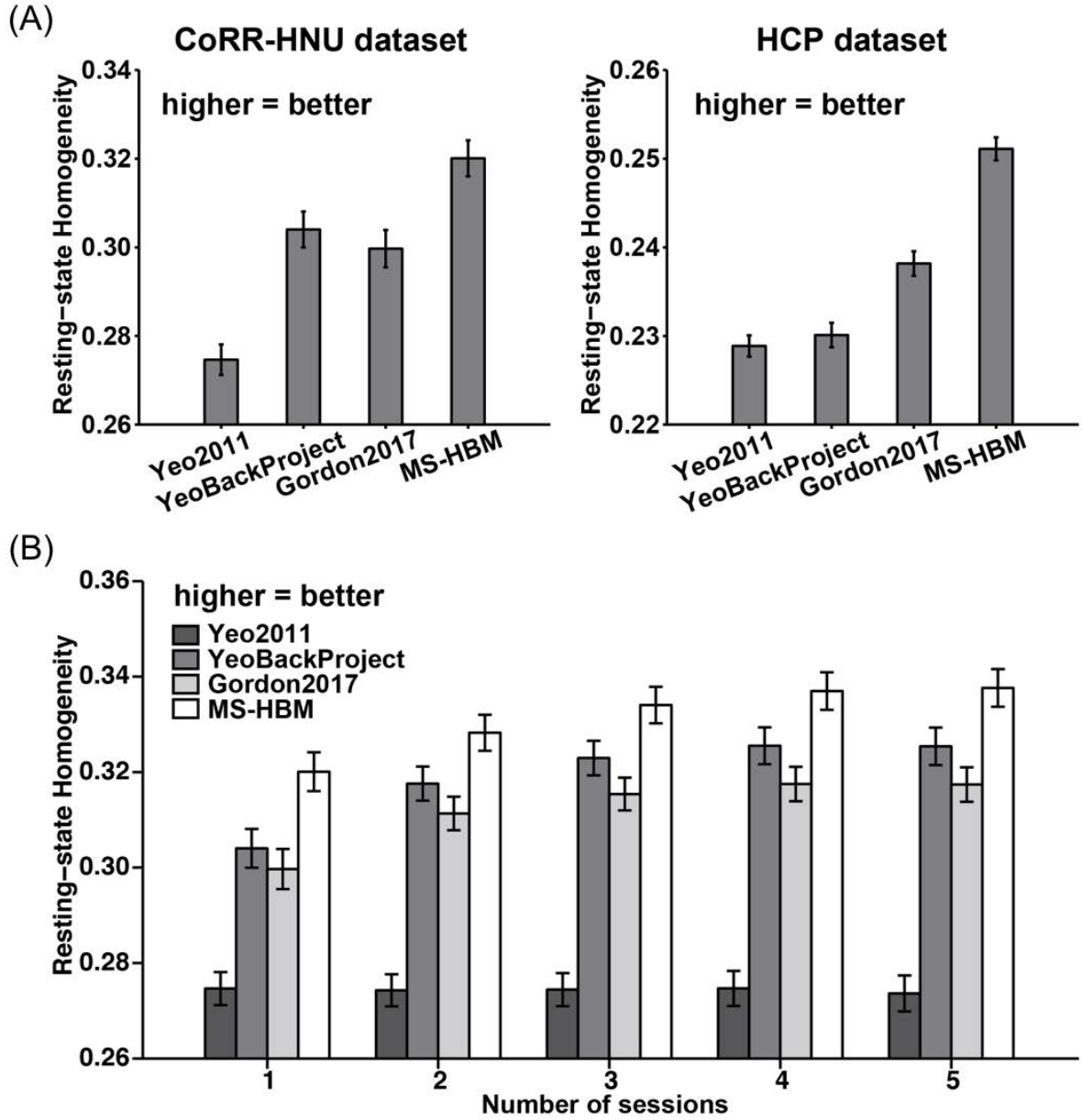

Figure 6. Resting-state homogeneity in the CoRR-HNU and GSP dataset. (A) 17-network individual-specific parcellations were estimated using one rs-fMRI session and resting-state homogeneity were computed on the remaining sessions for each subject from the CoRR-HNU and HCP dataset. (B) 17-network individual-specific parcellations were estimated using different number of rs-fMRI sessions and resting-state homogeneity were computed on the remaining sessions for each subject from the CoRR-HNU dataset. Using just one single fMRI sessions (10 min), the MS-HBM algorithm was able to match the homogeneity achieved with the Gordon2017 approach using five fMRI sessions (50 min).

Individual-specific networks generated by the MS-HBM exhibit lower task functional inhomogeneity than other approaches

Individual-specific parcellations were estimated using one resting-state fMRI session (15 min) from the HCP test set. The task inhomogeneity of the parcellations are shown in Figure 7. Compared with Yeo2011, YeoBackProject and Gordon2017, our approach achieved a modest average improvement of 0.54% ($p = 0.9$ for social, $p = 0.578$ for motor, $p < 5e\text{-}324$ for other 5 domains), 1.93% ($p < 5e\text{-}324$ for all domains) and 0.94% ($p < 5e\text{-}324$ for all domains) respectively. Interestingly, the Yeo2011 group-level parcellation performed as well as (or even better than) YeoBackProject and Gordon2017.

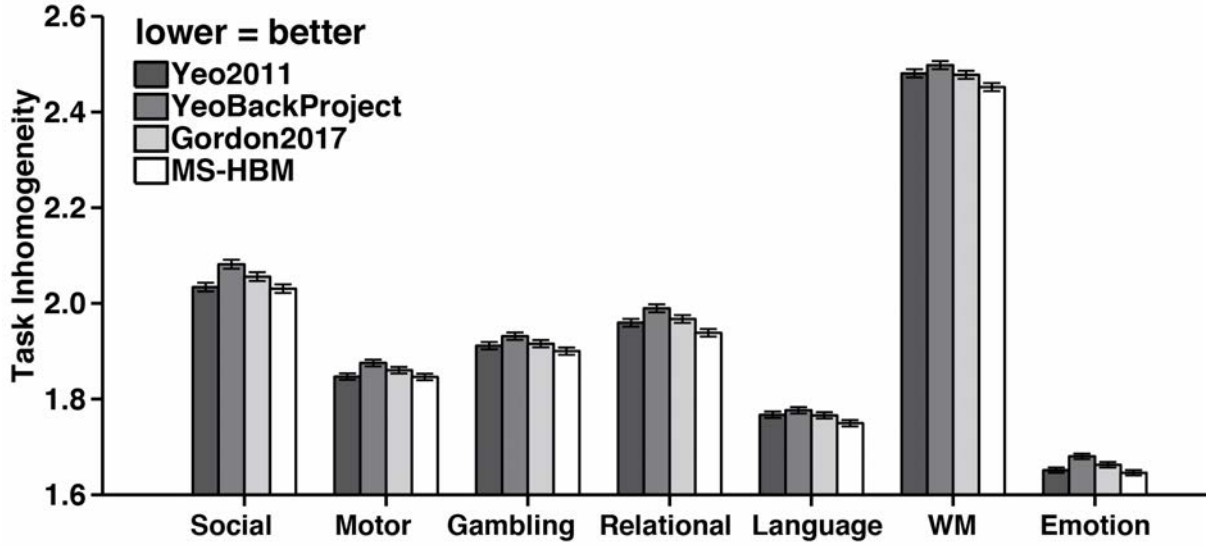

Figure 7. Task inhomogeneity of resting-state parcellations in the HCP dataset. 17-network individual-specific parcellations were estimated using one rs-fMRI session. Task inhomogeneity was then defined as the standard deviation of task activation within each network, and then averaged across all networks and contrasts within each behavioral domain. Lower value indicates better functional homogeneity. The MS-HBM individual-specific parcellations were more functional homogeneous than Yeo2011 ($p = 0.9056$ for social domain, $p = 0.578$ for motor domain, $p < 5e\text{-}324$ for other 5 domains), YeoBackProject ($p < 5e\text{-}324$ for all domains), Gordon2017 ($p < 5e\text{-}324$ for all domains).

Individual differences in cortical network parcellations are behaviorally relevant

Across all 58 behavioral measures, average prediction accuracy was $r = 0.084$ ($p < 4e\text{-}10$). While the accuracy might seem modest, they were comparable to results from the HCP MegaTrawl[1] (https://db.humanconnectome.org/megatrawl/). Figure 8 shows the prediction

[1] Of the 58 behavioral measures, 49 of them were also utilized in the HCP MegaTrawl. For the 300-dimensional group-ICA results, HCP MegaTrawl achieved an average accuracy of r

accuracy for 13 cognitive measures from the NIH toolbox. Average prediction accuracy was r = 0.15. The prediction accuracies for the remaining cognitive, emotion and personality measures are found in Figures S6 and S7. Interestingly, the prediction of emotional recognition (Figure S7) was poor with an average prediction accuracy of r = -0.036.

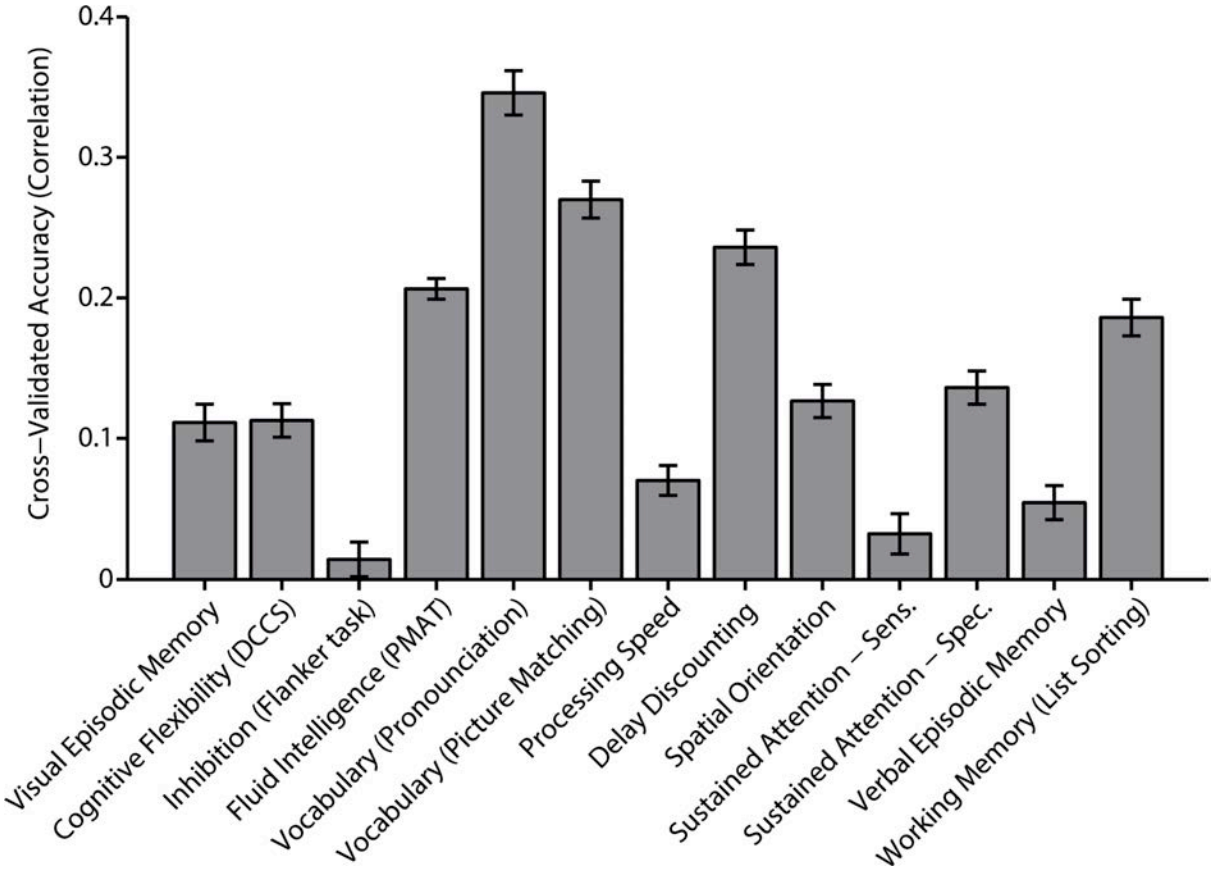

Figure 8. Prediction accuracy of 13 cognitive measures (NIH toolbox) based on inter-subject differences in the spatial arrangement of cortical networks. Average prediction accuracy is r = 0.15 for the 13 cognitive measures. Other behavioral measures are found in Figures S6 and S7.

= 0.059 (original data space), while kernel regression yielded an average accuracy of r = 0.091.

Discussion

We proposed a multi-session hierarchical Bayesian model (MS-HBM) that took into account inter-subject and intra-subject functional connectivity variability. The approach also differentiates between inter-subject functional connectivity variability and inter-subject network spatial probability. Across three multi-session datasets, we found that compared to association networks, sensory-motor networks exhibited lower inter-subject, but higher intra-subject functional connectivity variability. Sensory-motor networks also exhibited less inter-subject spatial variability than association networks. Furthermore, in both rs-fMRI and task-fMRI data, the MS-HBM individual-specific parcellations were more homogeneous than parcellations derived with three alternative approaches. Last but not least, we show that individual differences in the spatial arrangement of cortical networks can be used to predict individuals' cognition, emotion and personality.

Association networks exhibit more inter-subject variability than sensory-motor networks

The human association cortex is involved in higher cognitive functions (e.g., language, logical reasoning) that distinguish humans from other primates. Over the course of primate evolution, the human association cortex has undergone marked expansion, while the size of primary sensory cortices has largely stayed constant (Hill et al., 2010; Preuss 2011). This rapid expansion might result in massive functional and organizational differences between association and sensory cortices (Buckner & Krienen, 2013). For example, while information from the external milieu flows hierarchically from sensory to association cortex (Felleman and Van Essen, 1991), association regions are embedded within parallel, interdigitated networks (Goldman-Rakic 1988; Power et al., 2011; Yeo et al., 2011). As another example, a recent meta-analysis suggests that the association cortex contains complex zones ranging from being functionally specialized to functionally flexible, while sensory-motor regions are highly specialized (Yeo et al., 2015a).

Because the association cortex matures late during neurodevelopment (Hill et al., 2010; Buckner & Krienen, 2013), the prolonged exposure to environmental factors during a time of high neuroplasticity (Petanjek et al., 2011) might lead to greater individual differences in association cortical anatomy, function and connectivity. Indeed, anatomical studies have shown that early sensory-motor cortical areas (e.g., Area 17) exhibit less inter-subject spatial variability than association areas (e.g., Areas 44 and 45) after accounting for cortical folding patterns (Amunts et al., 1999; Amunts et al., 2000; Fischl et al., 2008; Yeo et

al., 2010a). As another example, the network of association regions supporting language processing is highly variable and can even be completely right lateralized in certain individuals (Geschwind & Levitsky 1968; Ojemann et al., 1989; Ojemann et al., 2003; Liu et al., 2009; Fedorenko et al., 2011; Frost & Goebel, 2012).

The hypothesis that association regions exhibit greater inter-subject functional connectivity variability than sensory-motor regions is strongly supported by recent rs-fMRI studies (Mueller et al., 2013; Chen et al., 2015; Laumann et al., 2015). One important methodological consideration is that previous studies assumed functional correspondence across subjects after macro-anatomical alignment (Mueller et al., 2013; Chen et al., 2015; Laumann et al., 2015). However, it is well-known that macro-anatomical alignment (or even functional alignment) is not sufficient to achieve perfect functional correspondence across subjects (Fischl et al 1999b; Yeo et al., 2010b; Robinson et al., 2014; Harrison et al., 2015; Langs et al., 2016; Glasser et al., 2016). Therefore, a portion of the inter-subject functional connectivity variability estimated from the previous studies might be simply a result of spatial misalignment of functional networks across subjects.

By explicitly differentiating between inter-subject spatial variability and inter-subject functional connectivity variability, we explored the possibility that for certain networks, inter-subject variability might be attributed to spatial variability, rather than functional connectivity variability. Interestingly, our results were largely in agreement with previous studies. Across three independent rs-fMRI datasets, compared with sensory-motor networks, association networks exhibited higher inter-subject functional connectivity (Figures 2, S1, S2) and spatial (Figures 3, S3, S4, 5). Among the association networks, the Salience/Ventral Attention network B was especially variable. The only networks with possible dissociation in inter-subject spatial and functional connectivity variability were the Limbic networks (A and B), which exhibited moderate inter-subject functional connectivity variability (Figure 2), but low inter-subject spatial variability (Figure 5).

Sensory-motor networks exhibit more intra-subject variability than association networks

An important criterion for a good biomarker is high test-retest reliability, which requires inter-subject differences to dominate intra-subject variability. Our current study suggests that intra-subject functional connectivity variability is substantially higher than inter-subject functional connectivity variability (Figure 2, S1, S2), consistent with many studies demonstrating that functional connectivity metrics exhibit moderate to high test-retest

reliability (Meindl et al., 2010; Van Dijk et al., 2010; Zuo et al., 2010; Wang et al., 2011; Guo et al., 2012; Zuo and Xing 2014; Chen et al., 2015; Xu et al., 2016).

While there have been many rs-fMRI test-retest studies, there are relatively few studies focusing on the spatial topography of intra-subject functional connectivity variability (Mueller et al., 2013; Chen et al., 2015; Laumann et al., 2015). Laumann and colleagues found that sensory-motor (visual, somatosensory, motor) regions exhibited high intra-subject functional connectivity variability, while association regions exhibited low intra-subject functional connectivity variability (Laumann et al., 2015). On the other hand, Mueller and colleagues found that low signal-to-noise regions (orbital frontal and temporal pole) exhibited high intra-subject variability, while portions of the default network exhibited low intra-subject variability (Mueller et al., 2013). Therefore, there were both agreements and discrepancies between the two studies[2]. Like before, it is worth noting that Mueller et al. (2013) assumed functional correspondence after macro-anatomical registration, while Laumann et al. (2015) utilized a subject-specific parcellation.

Unlike the previous works, our model differentiated between intra-subject and inter-subject functional connectivity variability, as well as inter-subject network spatial variability. Our results largely agreed with Laumann et al. (2015) in that sensory-motor networks exhibited high intra-subject variability, while association networks exhibited low intra-subject variability. Default networks (A and B) were the least variable, consistent with Mueller et al. (2013). These results were replicated across three datasets, although a particularly interesting difference is that Visual network B showed high intra-subject variability in the GSP dataset, but low or intermediate intra-subject variability in the CoRR-HNU and HCP datasets. This difference might be due to the fact that subjects were told to fixate on a cross in the CoRR-HNU and HCP datasets, while subjects were told to keep their eyes open (with no fixation cross) in the GSP dataset.

Intra-subject functional connectivity variability is probably the result of both measurement noise and participants' brain states. Measurement noise is influenced by acquisition protocol and scanner type, while participants' brain state might be influenced by factors, such as caffeine intake (Laumann et al., 2015; Poldrack et al., 2015), sleepiness (Yeo et al., 2015b; Laumann et al., 2016; Wang et al., 2016) and attention (Shine et al., 2016). However, measurement noise and participants' brain states are not independent factors. For

[2] Chen and colleagues (2015) utilized very different functional connectivity measures from the other two studies (Mueller et al., 2013; Laumann et al., 2015), so a direct comparison is difficult.

example, differences in acquisition protocols (e.g., absence or presence of a fixation cross) might lead to differences in participants' brain states, and thus, intra-subject variability. Indeed, it has been shown that many subjects fall asleep when instructed to close their eyes during the resting-state scan (Tagliazuchhi et al., 2014).

One approach to reduce intra-subject variability is to increase the acquisition time (Van Dijk et al., 2010; Xu et al., 2016). However, intra-subject functional connectivity variability is detectable even when concatenating many sessions of data (~100 minutes; Anderson et al., 2011; Laumann et al., 2015; Gordon et al., 2017c). Since intra-subject variability cannot be completely removed, a better parcellation strategy might theoretically be achieved by taking into account intra-subject variability.

Spatial configuration of individual-specific cortical parcellations is behaviorally meaningful

Given that inter-subject and intra-subject functional connectivity variability are different across functional brain networks, it is important for a parcellation strategy to distinguish between the two types of variability. For example, Somatomotor networks (A and B) exhibited low inter-subject, but high intra-subject, functional connectivity variability. A naïve algorithm might wrongly attribute differences in somatomotor connectivity between two subjects to inter-subject differences, rather than just inter-session noise.

The individual-specific parcellation approach in this paper modeled both inter-subject and intra-subject variability, allowing the identification of individual-specific functional networks that were highly reproducible within each subject, while also capturing variations across subjects (Figures 4, 5). Although all networks showed higher intra-subject reproducibility than inter-subject similarity, there were also differences across networks, with sensory-motor networks showing higher intra-subject reproducibility and higher inter-subject similarity than association networks (Figure 5).

Recent work has suggested that individual-specific functional networks exhibit unique topological features not observed in group-level networks (Harrison et al., 2015; Laumann et al., 2015; Glasser et al., 2016; Langs et al., 2016; Braga & Buckner, 2017; Gordon et al., 2017a; 2017b; 2017c). This is also clearly the case with individual-specific MS-HBM parcellations (Figure 4). While we have pointed out two examples (Default C and Control B networks), it is also obvious that many of these individual-specific parcellation features are replicable across multiple sessions.

A major unanswered question in the literature is whether individual differences in cortical parcellations are actually behaviorally meaningful. Here, kernel regression was

utilized to demonstrate that the spatial arrangement of individual-specific cortical networks can be used to predict cognition, emotion and personality in individual subjects. More specifically, kernel regression models the possibility that subjects with more similar parcellations exhibited similar behavior. Successful prediction suggests that inter-subject variation in the spatial configuration of cortical networks are strongly related to inter-subject variation in behavior.

Individual-specific MS-HBM parcellations are more homogeneous than other approaches during resting and task states

If an individual-specific parcellation is capturing the unique network-level organization of a subject's cerebral cortex, then vertices within each network should have similar resting-state time series, as well as similar activation amplitude for any given task contrast (Gordon et al., 2017c; Schaefer et al., in press). Across the CoRR-HNU and HCP datasets, individual-specific MS-HBM parcellations exhibited greater resting-state functional connectivity homogeneity than parcellations from three other approaches (Figure 6), suggesting that MS-HBM parcellations better capture the "intrinsic" organization of individuals' cerebral cortex. Importantly, model parameters (e.g., inter-subject and intra-subject variability) estimated from the GSP dataset could improve the estimation of individual-specific parcellations in the CoRR-HNU dataset (Figure 6A). This demonstration is important because estimates of inter-subject and intra-subject functional connectivity variability were similar, but not the same across datasets (Figures 2, S1, S2). Therefore, our results suggest that the MS-HBM approach can be used to parcellate individuals from new datasets (using the same preprocessing pipeline), without having to re-estimate the model parameters (e.g., inter-subject and intra-subject functional connectivity variability).

In the HCP dataset, individual-specific MS-HBM parcellations also exhibited greater task functional homogeneity than parcellations from three other approaches (Figure 7), suggesting that MS-HBM parcellations better capture the "extrinsic" organization of individuals' cerebral cortex. Given the strong link between task fMRI and resting-state fMRI (Smith et al., 2009; Mennes et al., 2010; Cole et al., 2014; Krienen et al., 2014; Bertolero et al., 2015; Yeo et al., 2015a; Tavor et al., 2016), this might not seem surprising. However, it is worth pointing that the group-level parcellation performed as well as, if not better than the two other individual-specific parcellation approaches (Figure 7). Furthermore, the MS-HBM approach only demonstrated (modest) improvements over the group-level parcellation in five of seven functional domains, while there was no statistical difference in the two remaining

two functional domains. One explanation is that the resting-state parcellations might be too coarse to capture the finer details of task activation. For example, the right-hand motor task preferentially activates the hand region of the left somatomotor cortex. However, the somatomotor network A is bilateral and covers the hand, foot and body regions of bilateral somatomotor cortex. As such, even if individual-specific somatomotor network A was highly accurate, the resulting task inhomogeneity might still be relatively high (see discussion on methodological considerations and future work).

MS-HBM approach works well with single-session rs-fMRI data

As discussed in a previous section, increasing the scan duration of resting-state fMRI can improve the reliability of functional connectivity measures (Van Dijk et al., 2010; Xu et al., 2016). While earlier studies have suggested that 5 to 12 minutes of resting-state scan might be sufficient to provide reliable measurements (Van Dijk et al., 2010; Birn et al., 2013), more recent studies have suggested the need for 25 to 30 minutes of data (Anderson et al., 2011; Laumann et al., 2015; Gordon et al., 2017c). However, it is important note that the amount of data necessary for reliable measurements depends on the functional connectivity measures being computed (Gordon et al., 2017c).

Consistent with previous studies, our experiments show that the quality of the individual-specific parcellations improves with more rs-fMRI data, although the improvements plateau after around 30 to 40 minutes of data (Figure 6B). Importantly, even though the MS-HBM was developed for multi-session rs-fMRI, the algorithm performed well even with single-session data. For example, the individual-specific MS-HBM parcellations estimated with one rs-fMRI session (10 minutes) exhibited comparable resting-state connectional homogeneity with parcellations estimated using a recent state-of-the-art approach with five times the amount of data (Gordon et al., 2017a, 2017b).

Methodological considerations and future work

Although the MS-HBM approach did not account for inter-site variability, we demonstrated that model parameters estimated from one site can generalize to another site with a different acquisition protocol (Figures 4, 5, 6). Given the increasing availability of multi-session rs-fMRI (Zuo et al., 2014; Holmes et al., 2015; Poldrack et al., 2015; Filevich et al., 2017; Gordon et al., 2017c), it might be possible to add another layer to the hierarchical model to account for inter-site variability, in addition to intra-subject and inter-subject variability. Furthermore, our experiments did not differentiate between rs-fMRI runs

collected within the same session versus rs-fMRI runs collected from different sessions. Another layer could again be inserted into the model to differentiate between intra-session and inter-session variability. However, we suspect diminishing returns.

By assuming individual-specific parcellations to be the same across sessions (Figure 1), the MS-HBM essentially treats inter-session differences as noise. The implication is that the individual-specific MS-HBM parcellations seek to capture stable, trait-like network organization in individuals. However, it is well-known that certain factors (e.g., caffeine intake and sleepiness) result in different brain states and thus functional network organization. Moreover, in longitudinal studies of certain populations, e.g., Alzheimer's Disease dementia, the goal is to detect neurological changes between consecutive sessions that are relatively far apart in time (Misra et al., 2009; Raj et al., 2015; Risacher et al., 2010; Zhang et al., 2016; Lindemer et al., 2017). To capture transient session-specific or longitudinal changes in brain network organization, the model could be modified to allow for possible spatial differences in individual-specific parcellations across sessions.

Here, we focused on parcellating the cerebral cortex into a small number of (less than twenty) networks. Each spatial (e.g., parietal) component of a network likely spans multiple cytoarchitectonically, functionally and connectionally distinct cortical areas (Kaas 1987; Felleman and Van Essen 1991; Amunts and Zilles 2015; Eickhoff et al., in press). It would be interesting to extend the MS-HBM to estimate a finer division of the cerebral cortex that might approximate classically defined cortical areas. The main challenge is that because of strong long-range functionally connectivity (Sepulcre et al., 2010), the MS-HBM will always result in spatially distributed networks even when estimating large number (e.g., hundreds) of networks. We are working on an additional spatial prior to ensure parcels are spatially localized, but not necessarily spatially connected (Glasser et al., 2016).

When preparing this manuscript, we became aware of related research suggesting that spatial variability can explain a significant portion of functional connectivity variability, and that spatial and functional connectivity variability explained similar behavioral variance (Bijsterbosch et al., biorxiv). However, given methodological differences in both estimating individual-specific brain networks and predicting behavior from our current work, we believe that it would still be worthwhile investigating whether inter-subject network spatial variability and inter-subject functional connectivity variability can be combined to improve the prediction of individuals' behavior.

Conclusions

We developed a multi-session hierarchical Bayesian model (MS-HBM) that differentiated between inter-subject and intra-subject variability when estimating individual-specific cortical network parcellations. Across three datasets, sensory-motor networks exhibited lower inter-subject, but higher intra-subject functional connectivity variability than association networks. Sensory-motor networks were also more spatially variable across subjects than association networks. The individual-specific MS-HBM parcellations were highly reproducible within individuals while capturing network variations across subjects. Using a single rs-fMRI session (10 min), our approach yielded parcellations comparable to those estimated by a recent template matching algorithm using five rs-fMRI sessions (50 min). Finally, we showed that inter-subject variation in the spatial configuration of cortical networks are strongly related to inter-subject variation in behavior.

Acknowledgement

This work was supported by Singapore MOE Tier 2 (MOE2014-T2-2-016), NUS Strategic Research (DPRT/944/09/14), NUS SOM Aspiration Fund (R185000271720), Singapore NMRC (CBRG/0088/2015), NUS YIA, the Singapore National Research Foundation (NRF) Fellowship (Class of 2017). AJH was supported by the National Institute of Mental Health (573 K01MH099232). XNZ was supported by the National Basic Research (973) Program (2015CB351702), the Natural Science Foundation of China (81471740, 81220108014), and the Beijing Municipal Science and Tech Commission (Z161100002616023, Z171100000117012). Our research also utilized resources provided by the Center for Functional Neuroimaging Technologies, P41EB015896 and instruments supported by 1S10RR023401, 1S10RR019307, and 1S10RR023043 from the Athinoula A. Martinos Center for Biomedical Imaging at the Massachusetts General Hospital. Data were also provided by the Brain Genomics Superstruct Project of Harvard University and the Massachusetts General Hospital (Principal Investigators: Randy Buckner, Joshua Roffman, and Jordan Smoller), with support from the Center for Brain Science Neuroinformatics Research Group, the Athinoula A. Martinos Center for Biomedical Imaging, and the Center for Human Genetic Research. Twenty individual investigators at Harvard and MGH generously contributed data to the overall project. Data were also provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support

the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

References

Amunts K, Malikovic A, Mohlberg H, Schormann T, Zilles K. 2000. Brodmann's Areas 17 and 18 Brought into Stereotaxic Space—Where and How Variable? NeuroImage. 11:66–84.

Amunts K, Schleicher A, B rgel U, Mohlberg H, Uylings HBM, Zilles K. 1999. Broca's region revisited: Cytoarchitecture and intersubject variability. The Journal of Comparative Neurology. 412:319–341.

Amunts K, Zilles K. 2015. Architectonic mapping of the human brain beyond Brodmann. Neuron. 88:1086–1107.

Anderson JS, Ferguson MA, Lopez-Larson M, Yurgelun-Todd D. 2011. Reproducibility of Single-Subject Functional Connectivity Measurements. American Journal of Neuroradiology. 32:548–555.

Barch DM, Burgess GC, Harms MP, Petersen SE, Schlaggar BL, Corbetta M, Glasser MF, Curtiss S, Dixit S, Feldt C, Nolan D, Bryant E, Hartley T, Footer O, Bjork JM, Poldrack R, Smith S, Johansen-Berg H, Snyder AZ, Van Essen DC. 2013. Function in the human connectome: Task-fMRI and individual differences in behavior. Neuroimage. 80:169–189.

Beckmann C, Mackay C, Filippini N, Smith S. 2009. Group comparison of resting-state fMRI data using multi-subject ICA and dual regression. 15th Annual Meeting of Organization for Human Brain Mapping. poster 441 SU-AM.

Beckmann CF, DeLuca M, Devlin JT, Smith SM. 2005. Investigations into resting-state connectivity using independent component analysis. Philos Trans R Soc B Biol Sci. 360:1001–1013.

Bellec P, Rosa-Neto P, Lyttelton OC, Benali H, Evans AC. 2010. Multi-level bootstrap analysis of stable clusters in resting-state fMRI. Neuroimage. 51:1126–1139.

Bertolero MA, Yeo BTT, D'Esposito M. 2015. The modular and integrative functional architecture of the human brain. Proc Natl Acad Sci. 112:E6798–E6807.

Biswal B, Zerrin Yetkin F, Haughton VM, Hyde JS. 1995. Functional Connectivity in the Motor Cortex of Resting Human Brain Using Echo-Planar MRI. Magn Reson Med. 34:537–541.

Bijsterbosch JD, Woolrich MW, Glasser MF, Robinson EC, Beckmann CF, Van Essen DC, Harrison SJ and Smith SM. 2017. The relationship between spatial configuration and functional connectivity of brain regions. bioRxiv. doi: https://www.biorxiv.org/content/early/2017/10/27/210195

Bouckaert RR, Frank E. 2004. Evaluating the Replicability of Significance Tests for Comparing Learning Algorithms. In: Information Processing in Medical Imaging. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg. p. 3–12.

Braga RM, Buckner RL. 2017. Parallel interdigitated distributed networks within the individual estimated by intrinsic functional connectivity. Neuron. 95:457-471.

Buckner RL, Krienen FM, Castellanos A, Diaz JC, Yeo BTT. 2011. The organization of the human cerebellum estimated by intrinsic functional connectivity. J Neurophysiol. 106:2322–2345.

Buckner RL, Krienen FM, Yeo BTT. 2013. Opportunities and limitations of intrinsic functional connectivity MRI. Nat Neurosci. 16:832–837.

Buckner RL, Krienen FM. 2013. The evolution of distributed association networks in the human brain. Trends Cogn Sci. 17:648–665.

Burgess GC, Kandala S, Nolan D, Laumann TO, Power JD, Adeyemo B, Harms MP, Petersen SE, Barch DM. 2016. Evaluation of Denoising Strategies to Address Motion-Correlated Artifacts in Resting-State Functional Magnetic Resonance Imaging Data

from the Human Connectome Project. Brain Connectivity. 6:669–680.
Calhoun VD, Adali T. 2012. Multisubject Independent Component Analysis of fMRI: A Decade of Intrinsic Networks, Default Mode, and Neurodiagnostic Discovery. IEEE Rev Biomed Eng. 5:60–73.
Calhoun VD, Kiehl KA, Pearlson GD. 2008. Modulation of temporally coherent brain networks estimated using ICA at rest and during cognitive tasks. Hum Brain Mapp. 29:828–838.
Calhoun VD, Liu J, Adalı T. 2009. A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. Neuroimage. 45:S163-S172.
Chen B, Xu T, Zhou C, Wang L, Yang N, Wang Z, Dong H-M, Yang Z, Zang Y-F, Zuo X-N, Weng X-C. 2015. Individual Variability and Test-Retest Reliability Revealed by Ten Repeated Resting-State Brain Scans over One Month. PLoS ONE. 10:e0144963–21.
Chong M, Bhushan C, Joshi AA, Choi S, Haldar JP, Shattuck DW, Spreng RN, Leahy RM. 2017. Individual parcellation of resting fMRI with a group functional connectivity prior. NeuroImage. 156:87–100.
Cole MW, Bassett DS, Power JD, Braver TS, Petersen SE. 2014. Intrinsic and task-evoked network architectures of the human brain. Neuron. 83:238–251.
Dale AM, Fischl B, Sereno MI. 1999. Cortical surface-based analysis. I. Segmen- tation and surface reconstruction. Neuroimage. 9:179–194.
Damoiseaux JS, Rombouts SARB, Barkhof F, Scheltens P, Stam CJ, Smith SM, Beckmann CF. 2006. Consistent resting-state networks across healthy subjects. Proc Natl Acad Sci U S A. 103:13848–13853.
Dosenbach NUF, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RAT, Fox MD, Snyder AZ, Vincent JL, Raichle ME, others. 2007. Distinct brain networks for adaptive and stable task control in humans. Proc Natl Acad Sci. 104:11073–11078.
Eickhoff SB, Constable RT, Yeo BTT. 2017. Topographic organization of the cerebral cortex and brain cartography. NeuroImage. In press
Fedorenko E, Behr MK, Kanwisher N. 2011. Functional specificity for high-level linguistic processing in the human brain. Proceedings of the National Academy of Sciences. 108:16428–16433.
Felleman DJ, Van Essen DC. 1991. Distributed Hierarchical Processing in the Primate Cerebral Cortex. Cereb Cortex. 1:1–47.
Filevich E, Lisofsky N, Becker M, Butler O, Lochstet M, Martensson J, Wenger E, Lindenberger U, Kühn S. 2017. Day2day: investigating daily variability of magnetic resonance imaging measures over half a year. BMC Neuroscience. 18:161.
Filippini N, MacIntosh BJ, Hough MG, Goodwin GM, Frisoni GB, Smith SM, Matthews PM, Beckmann CF, Mackay CE. 2009. Distinct patterns of brain activity in young carriers of the APOE-ε4 allele. Proceedings of the National Academy of Sciences. 106:7209-7214.
Finn ES, Shen X, Scheinost D, Rosenberg MD, Huang J, Chun MM, Papademetris X, Constable RT. 2015. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat Neurosci. 18:1664-1671.
Fischl B, Liu A, Dale AM. 2001. Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. IEEE Trans Med Imaging. 20:70–80.
Fischl B, Rajendran N, Busa E, Augustinack J, Hinds O, Yeo BTT, Mohlberg H, Amunts K, Zilles K. 2008. Cortical Folding Patterns and Predicting Cytoarchitecture. Cereb Cortex. 18:1973–1980.
Fischl B, Sereno MI, Dale AM. 1999a. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. Neuroimage. 9:195–207.
Fischl B, Sereno MI, Tootell RBH, Dale AM. 1999b. High-resolution intersubject averaging

and a coordinate system for the cortical surface. Hum Brain Mapp. 8:272–284.
Fox MD, Corbetta M, Snyder AZ, Vincent JL, Raichle ME. 2006. Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. Proc Natl Acad Sci. 103:10046–10051.
Fox MD, Raichle ME. 2007. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. Nat Rev Neurosci. 8:700–711.
Frost MA, Goebel R. 2012. Measuring structural–functional correspondence: Spatial variability of specialised brain regions after macro-anatomical alignment. NeuroImage. 59:1369–1381.
Ge T, Holmes AJ, Buckner RL, Smoller JW, Sabuncu MR. 2017. Heritability analysis with repeat measurements and its application to resting-state functional connectivity. Proceedings of the National Academy of Sciences. 114:5521–5526.
Geschwind N, Levitsky W. 1968. Human Brain: Left-Right Asymmetries in Temporal Speech Region. Science. 161:186–187.
Glahn DC, Winkler AM, Kochunov P, Almasy L, Duggirala R, Carless MA, Curran JC, Olvera RL, Laird AR, Smith SM, Beckmann CF, Fox PT, Blangero J. 2010. Genetic control over the resting brain. Proc Natl Acad Sci U S A. 107:1223–1228.
Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, Uğurbil K, Andersson J, Beckmann CF, Jenkinson M, Smith SM, Van Essen DC. 2016. A multi-modal parcellation of human cerebral cortex. Nature Publishing Group. 536:171–178.
Glasser MF, Sotiropoulos SN, Wilson JA, Coalson T, Fischl B, Andersson JL, Xu J, Jbabdi S, Webster M, Polimeni JR, Van Essen DC, Jenkinson M. 2013. The minimal preprocessing pipelines for the Human Connectome Project. NeuroImage. 80:105-124.
Goldman-Rakic PS. 1988. Topography of Cognition: Parallel Distributed Networks in Primate Association Cortex. Annu Rev Neurosci. 11:137–156.
Gordon EM, Laumann TO, Adeyemo B, Gilmore AW, Nelson SM, Dosenbach NUF, Petersen SE. 2017b. Individual-specific features of brain systems identified with resting state functional correlations. NeuroImage. 146:918–939.
Gordon EM, Laumann TO, Adeyemo B, Huckins JF, Kelley WM, Petersen SE. 2016. Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. Cereb Cortex. 26:288–303.
Gordon EM, Laumann TO, Adeyemo B, Huckins JF, Kelley WM, Petersen SE. 2015. Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. Cereb Cortex. 26:288–303.
Gordon EM, Laumann TO, Adeyemo B, Petersen SE. 2017a. Individual Variability of the System-Level Organization of the Human Brain. Cereb Cortex. 27:386-399.
Gordon EM, Laumann TO, Gilmore AW, Newbold DJ, Greene DJ, Berg JJ, Ortega M, Hoyt-Drazen C, Gratton C, Sun H, Hampton JM, Coalson RS, Nguyen AL, McDermott KB, Shimony JS, Snyder AZ, Schlaggar BL, Petersen SE, Nelson SM, Dosenbach NUF. 2017c. Precision Functional Mapping of Individual Human Brains. Neuron. 1–25.
Greicius MD, Krasnow B, Reiss AL, Menon V. 2003. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc Natl Acad Sci. 100:253–258.
Greve DN, Fischl B. 2009. Accurate and robust brain image alignment using boundary-based registration. Neuroimage. 48:63–72.
Griffanti L, Salimi-Khorshidi G, Beckmann CF, Auerbach EJ, Douaud G, Sexton CE, Zsoldos E, Ebmeier KP, Filippini N, Mackay CE, Moeller S, Xu J, Yacoub E, Baselli G, Ugurbil K, Miller KL, Smith SM. 2014. ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. Neuroimage. 95:232–247.
Guo CC, Kurth F, Zhou J, Mayer EA, Eickhoff SB, Kramer JH, Seeley WW. 2012. One-year

test–retest reliability of intrinsic connectivity network fMRI in older adults. NeuroImage. 61:1471–1483.
Hacker CD, Laumann TO, Szrama NP, Baldassarre A, Snyder AZ, Leuthardt EC, Corbetta M. 2013. Resting state network estimation in individual subjects. Neuroimage. 82:616–633.
Hampson M, Driesen NR, Skudlarski P, Gore JC, Constable RT. 2006. Brain connectivity related to working memory performance. J Neurosci. 26:13338-13343.
Harrison SJ, Woolrich MW, Robinson EC, Glasser MF, Beckmann CF, Jenkinson M, Smith SM. 2015. Large-scale Probabilistic Functional Modes from resting state fMRI. Neuroimage. 109:217–231.
Hawrylycz M, Miller JA, Menon V, Feng D, Dolbeare T, Guillozet-Bongaarts AL, Jegga AG, Aronow BJ, Lee C-K, Bernard A, others. 2015. Canonical genetic signatures of the adult human brain. Nat Neurosci. 18:1832–1844.
Hill J, Inder T, Neil J, Dierker D, Harwell J, Van Essen D. 2010. Similar patterns of cortical expansion during human development and evolution. Proceedings of the National Academy of Sciences. 107:13135–13140.
Holmes AJ, Hollinshead MO, O'Keefe TM, Petrov VI, Fariello GR, Wald LL, Fischl B, Rosen BR, Mair RW, Roffman JL, Smoller JW, Buckner RL. 2015. Brain Genomics Superstruct Project initial data release with structural, functional, and behavioral measures. Sci data. 2:150031.
Jenkinson M, Bannister P, Brady M, Smith S. 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage. 17:825–841.
Kaas JH. 1987. The Organization of Neocortex in Mammals: Implications for Theories of Brain Function. Annual Review of Psychology. 38:129–151.
Krienen FM, Yeo BTT, Buckner RL. 2014. Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture. Phil Trans R Soc B. 369:20130526.
Krienen FM, Yeo BTT, Ge T, Buckner RL, Sherwood CC. 2016. Transcriptional profiles of supragranular-enriched genes associate with corticocortical network architecture in the human brain. Proc Natl Acad Sci. 113:E469–E478.
Kwong KK, Belliveau JW, Chesler DA, Goldberg IE, Weisskoff RM, Poncelet BP, Kennedy DN, Hoppel BE, Cohen MS, Turner R. 1992. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. Proc Natl Acad Sci. 89:5675–5679.
Langs G, Wang D, Golland P, Mueller S, Pan R, Sabuncu MR, Sun W, Li K, Liu H. 2016. Identifying Shared Brain Networks in Individuals by Decoupling Functional and Anatomical Variability. Cereb Cortex. 26:4004–4014.
Lashkari D, Vul E, Kanwisher N, Golland P. 2010. Discovering structure in the space of fMRI selectivity profiles. NeuroImage. 50:1085–1098.
Laumann TO, Gordon EM, Adeyemo B, Snyder AZ, Joo SJ, Chen M-Y, Gilmore AW, McDermott KB, Nelson SM, Dosenbach NUF, others. 2015. Functional system and areal organization of a highly sampled individual human brain. Neuron. 87:657–670.
Laumann TO, Snyder AZ, Mitra A, Gordon EM, Gratton C, Adeyemo B, Gilmore AW, Nelson SM, Berg JJ, Greene DJ, McCarthy JE, Tagliazucchi E, Laufs H, Schlaggar BL, Dosenbach NUF, Petersen SE. 2016. On the Stability of BOLD fMRI Correlations. Cereb Cortex.
Lee MH, Hacker CD, Snyder AZ, Corbetta M, Zhang D, Leuthardt EC, Shimony JS. 2012. Clustering of resting state networks. PLoS One. 7:e40370.
Lindemer ER, Greve DN, Fischl BR, Augustinack JC, Salat DH. 2017. Regional staging of white matter signal abnormalities in aging and Alzheimer's disease. NeuroImage:

Clinical. 14:156–165.
Liu H, Stufflebeam SM, Sepulcre J, Hedden T, Buckner RL. 2009. Evidence from intrinsic activity that asymmetry of the human brain is controlled by multiple factors. Proceedings of the National Academy of Sciences. 106:20499–20503.
Margulies DS, Kelly AMC, Uddin LQ, Biswal BB, Castellanos FX, Milham MP. 2007. Mapping the functional connectivity of anterior cingulate cortex. Neuroimage. 37:579–588.
Meindl T, Teipel S, Elmouden R, Mueller S, Koch W, Dietrich O, Coates U, Reiser M, Glaser C. 2010. Test-retest reproducibility of the default-mode network in healthy individuals. Hum Brain Mapp. 31:237–246.
Misra C, Fan Y, Davatzikos C. 2009. Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: Results from ADNI. NeuroImage. 44:1415–1422.
Mueller S, Wang D, Fox MD, Yeo BTT, Sepulcre J, Sabuncu MR, Shafee R, Lu J, Liu H. 2013. Individual variability in functional connectivity architecture of the human brain. Neuron. 77:586–595.
Murphy KP. 2012. Machine Learning. MIT Press.
Nadeau C and Bengio Y. 2000. Inference for the generalization error. In Advances in neural information processing systems, pp. 307-313.
Ogawa S, Tank DW, Menon R, Ellermann JM, Kim SG, Merkle H, Ugurbil K. 1992. Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. Proc Natl Acad Sci. 89:5951–5955.
Ojemann G, Ojemann J, Lettich EREEGT, Berger M. 1989. Cortical language localization in left, dominant hemisphere: an electrical stimulation mapping investigation in 117 patients. Journal of neurosurgery. 71:316-326.
Ojemann GA. 2003. The neurobiology of language and verbal memory: observations from awake neurosurgery. International Journal of Psychophysiology. 48:141–146.
Petanjek Z, Judas M, Simic G, Rasin MR, Uylings HBM, Rakic P, Kostovic I. 2011. Extraordinary neoteny of synaptic spines in the human prefrontal cortex. Proceedings of the National Academy of Sciences. 108:13281–13286.
Poldrack RA, Laumann TO, Koyejo O, Gregory B, Hover A, Chen M-Y, Gorgolewski KJ, Luci J, Joo SJ, Boyd RL, Hunicke-Smith S, Simpson ZB, Caven T, Sochat V, Shine JM, Gordon E, Snyder AZ, Adeyemo B, Petersen SE, Glahn DC, Reese Mckay D, Curran JE, Göring HHH, Carless MA, Blangero J, Dougherty R, Leemans A, Handwerker DA, Frick L, Marcotte EM, Mumford JA. 2015. Long-term neural and physiological phenotyping of a single human. Nature Communications. 6:8885.
Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, Petersen SE. 2011. Functional network organization of the human brain. Neuron. 72:665–678.
Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE. 2014. Methods to detect, characterize, and remove motion artifact in resting state fMRI. Neuroimage. 84:320–341.
Preuss TM. 2011. The human brain: rewired and running hot. Annals of the New York Academy of Sciences. 1225:E182–E191.
Raj A, LoCastro E, Kuceyeski A, Tosun D, Relkin N, Weiner M. 2015. Network Diffusion Model of Progression Predicts Longitudinal Patterns of Atrophy and Metabolism in Alzheimer's Disease. Cell Reports. 10:359–369.
Richiardi J, Altmann A, Milazzo A-C, Chang C, Chakravarty MM, Banaschewski T, Barker GJ, Bokde ALW, Bromberg U, Büchel C, others. 2015. Correlated gene expression supports synchronous activity in brain networks. Science (80-). 348:1241–1244.

Risacher SL, Shen L, West JD, Kim S, McDonald BC, Beckett LA, Harvey DJ, Jack CR Jr., Weiner MW, Saykin AJ. 2010. Longitudinal MRI atrophy biomarkers: Relationship to conversion in the ADNI cohort. Neurobiology of Aging. 31:1401–1418.

Robinson EC, Jbabdi S, Glasser MF, Andersson J, Burgess GC, Harms MP, Smith SM, Van Essen DC, Jenkinson M. 2014. MSM: A new flexible framework for Multimodal Surface Matching. NeuroImage. 100:414–426.

Rosenberg MD, Finn ES, Scheinost D, Papademetris X, Shen X, Constable RT, Chun MM. 2016. A neuromarker of sustained attention from whole-brain functional connectivity. Nat Neurosci. 19:165–171.

Salimi-Khorshidi G, Douaud G, Beckmann CF, Glasser MF, Griffanti L, Smith SM. 2014. Automatic denoising of functional MRI Data: Combining independent component analysis and hierarchical fusion of classifiers. Neuroimage. 90:449-468.

Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo X-N, Holmes AJ, Eickhoff SB, Yeo BTT. 2017. Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. Cereb Cortex. 1–20.

Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, Reiss AL, Greicius MD. 2007. Dissociable intrinsic connectivity networks for salience processing and executive control. J Neurosci. 27:2349–2356.

Ségonne F, Pacheco J, Fischl B. 2007. Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. IEEE Trans Med Imaging. 26:518–529.

Sepulcre J, Liu H, Talukdar T, Martincorena I, Yeo BTT, Buckner RL. 2010. The Organization of Local and Distant Functional Connectivity in the Human Brain. PLoS Comput Biol. 6:e1000808–e1000815.

Shine JM, Koyejo O, Poldrack RA. 2016. Temporal metastates are associated with differential patterns of time-resolved connectivity, network topology, and attention. Proceedings of the National Academy of Sciences. 113:9888–9891.

Siegel JS, Mitra A, Laumann TO, Seitzman BA, Raichle M, Corbetta M, Snyder AZ. 2017. Data Quality Influences Observed Links Between Functional Connectivity and Behavior. Cereb Cortex. 27:4492–4502

Smith SM, Beckmann CF, Andersson J, Auerbach EJ, Bijsterbosch J, Douaud G, Duff E, Feinberg DA, Griffanti L, Harms MP, Kelly M, Laumann T, Miller KL, Moeller S, Petersen S, Power J, Salimi-Khorshidi G, Snyder AZ, Vu AT, Woolrich MW, Xu J, Yacoub E, Uğurbil K, Van Essen DC, Glasser MF. 2013. Resting-state fMRI in the Human Connectome Project. Neuroimage. 80:144–168.

Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, Filippini N, Watkins KE, Toro R, Laird AR, others. 2009. Correspondence of the brain's functional architecture during activation and rest. Proc Natl Acad Sci. 106:13040–13045.

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney D, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM. 2004. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage. 23:S208–S219.

Smith SM, Nichols TE, Vidaurre D, Winkler AM, Behrens TE, Glasser MF, Ugurbil K, Barch DM, Van Essen DC, Miller KL. 2015. A positive-negative mode of population covariation links brain connectivity, demographics and behavior. Nat Neurosci. 18:1565-1567.

Tagliazucchi E, Laufs H. 2014. Decoding Wakefulness Levels from Typical fMRI Resting-State Data Reveals Reliable Drifts between Wakefulness and Sleep. Neuron. 82:695–708.

Tavor I, Jones OP, Mars RB, Smith SM, Behrens TE, Jbabdi S. 2016. Task-free MRI predicts individual differences in brain activity during task performance. Science (80-).

352:216–220.
van den Heuvel M, Mandl R, Hulshoff Pol H. 2008. Normalized Cut Group Clustering of Resting-State fMRI Data. PLoS One. 3:e2001.
van den Heuvel MP, Stam CJ, Kahn RS, Hulshoff Pol HE. 2009. Efficiency of functional brain networks and intellectual performance. J Neurosci. 29:7619–7624
Van Dijk KRA, Hedden T, Venkataraman A, Evans KC, Lazar SW, Buckner RL. 2010. Intrinsic Functional Connectivity As a Tool For Human Connectomics: Theory, Properties, and Optimization. Journal of Neurophysiology. 103:297–321.
Van Essen DC, Glasser MF, Dierker DL, Harwell J, Coalson T. 2012a. Parcellations and hemispheric asymmetries of human cerebral cortex analyzed on surface-based atlases. Cereb Cortex. 22:2241-62.
Van Essen DC, Ugurbil K, Auerbach E, Barch D, Behrens TEJ, Bucholz R, Chang A, Chen L, Corbetta M, Curtiss SW, Della Penna S, Feinberg D, Glasser MF, Harel N, Heath AC, Larson-Prior L, Marcus D, Michalareas G, Moeller S, Oostenveld R, Petersen SE, Prior F, Schlaggar BL, Smith SM, Snyder AZ, Xu J, Yacoub E. 2012b. The Human Connectome Project: A data acquisition perspective. Neuroimage. 62:2222–2231.
Varoquaux G, Gramfort A, Pedregosa F, Michel V, Thirion B. 2011. Multi-subject Dictionary Learning to Segment an Atlas of Brain Spontaneous Activity. In: Information Processing in Medical Imaging. p. 562–573.
Vincent JL, Snyder AZ, Fox MD, Shannon BJ, Andrews JR, Raichle ME, Buckner RL. 2006. Coherent spontaneous activity identifies a hippocampal-parietal memory network. J Neurophysiol. 96:3517–3531.
Wang C, Ong JL, Patanaik A, Zhou J, Chee MWL. 2016. Spontaneous eyelid closures link vigilance fluctuation with fMRI dynamic connectivity states. Proceedings of the National Academy of Sciences. 113:9653–9658.
Wang D, Buckner RL, Fox MD, Holt DJ, Holmes AJ, Stoecklein S, Langs G, Pan R, Qian T, Li K, Baker JT. 2015. Parcellating cortical functional networks in individuals. Nat Neurosci, 18:1853–1860.
Wang JH, Zuo XN, Gohel S, Milham MP, Biswal BB, He Y. 2011. Graph Theoretical Analysis of Functional Brain Networks: Test-Retest Evaluation on Short- and Long-Term Resting-State Functional MRI Data. PLoS ONE. 6:e21976–22.
Parcellating an Individual Subject's Cortical and Subcortical Brain Structures Using Snowball Sampling of Resting-State Correlations. 2013. Parcellating an Individual Subject's Cortical and Subcortical Brain Structures Using Snowball Sampling of Resting-State Correlations. 24:2036–2054.
Xu T, Opitz A, Craddock RC, Wright MJ, Zuo X-N, Milham MP. 2016. Assessing Variations in Areal Organization for the Intrinsic Brain: From Fingerprints to Reliability. Cereb Cortex. 26:4192–4211.
Yang Z, Zuo X-N, McMahon KL, Craddock RC, Kelly C, de Zubicaray GI, Hickie I, Bandettini PA, Castellanos FX, Milham MP, others. 2016. Genetic and Environmental Contributions to Functional Connectivity Architecture of the Human Brain. Cereb Cortex. 26:2341–2352.
Yeo BTT, Krienen FM, Eickhoff SB, Yaakub SN, Fox PT, Buckner RL, Asplund CL, Chee MWL. 2015a. Functional Specialization and Flexibility in Human Association Cortex. Cereb Cortex. 25:3654–3672.
Yeo BTT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, Fischl B, Liu H, Buckner RL. 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol. 106:1125–1165.
Yeo BTT, Sabuncu MR, Vercauteren T, Ayache N, Fischl B, Golland P. 2010a. Spherical

Demons: Fast Diffeomorphic Landmark-Free Surface Registration. IEEE Trans Med Imaging. 29:650–668.
Yeo BTT, Sabuncu MR, Vercauteren T, Holt DJ, Amunts K, Zilles K, Golland P, Fischl B. 2010b. Learning Task-Optimal Registration Cost Functions for Localizing Cytoarchitecture and Function in the Cerebral Cortex. IEEE Trans Med Imaging. 29:1424–1441.
Yeo BTT, Tandi J, Chee MWL. 2015b. Functional connectivity during rested wakefulness predicts vulnerability to sleep deprivation. Neuroimage. 111:147–158.
Zhang X, Mormino EC, Sun N, Sperling RA, Sabuncu MR, Yeo BTT, the Alzheimer's Disease Neuroimaging Initiative. 2016. Bayesian model reveals latent atrophy factors with dissociable cognitive trajectories in Alzheimer's disease. Proceedings of the National Academy of Sciences. 113:E6535–E6544.
Zuo XN, Anderson JS, Bellec P, Birn RM, Biswal BB, Blautzik J, Breitner JC, Buckner RL, Calhoun VD, Castellanos FX, Chen A. 2014. An open science resource for establishing reliability and reproducibility in functional connectomics. Sci data. 1:140049.
Zuo XN, Kelly C, Adelstein JS, Klein DF, Castellanos FX, Milham MP. 2010. Reliable intrinsic connectivity networks: Test–retest evaluation using ICA and dual regression approach. NeuroImage. 49:2163–2177.
Zuo XN, Xing XX. 2014. Test-retest reliabilities of resting-state FMRI measurements in human brain functional connectomics: A systems neuroscience perspective. Neuroscience and Biobehavioral Reviews. 45:100–118.

Controlling for Intra-Subject and Inter-Subject Variability in Individual-Specific Cortical Network Parcellations

Supplemental Material

This supplemental material is divided into *Supplemental Methods* and *Supplemental Results* to complement the Methods and Results sections in the main text, respectively.

Supplementary Methods

This section provides additional mathematical and implementation details of the multi-session hierarchical Bayesian model (MS-HBM). Section S1 provides mathematical details about the generative model. Section S2 describe the algorithms for estimating group-level priors and deriving the individual-specific parcellations. Section S3 provides details on how "free" parameters of the model are set.

S1. Mathematical model

In this section, we describe our model for individual-level parcellation of the cerebral cortex. We assume a common surface coordinate system, where the cerebral cortex is represented by left and right hemisphere spherical meshes such as FreeSurfer fsaverage surface meshes. Each mesh consists of a collection of vertices and edges connecting neighboring vertices into triangles (https://en.wikipedia.org/wiki/Triangle_mesh).

Let N denote the total number of vertices, T denote the number of resting-state fMRI (rs-fMRI) sessions, S denote the number of subjects, L denote the number of networks, and $\mathcal{N}_n$ denote the neighboring vertices of vertex n (as defined by the cortical mesh). For each subject s and session t, there is a preprocessed rs-fMRI time course associated with each vertex n. For each subject s, there is an unknown parcellation label l_n^s at vertex n. Note that the parcellation label is assumed to be the same across sessions (hence there is no index on the session). In this work, we use $1{:}S$ to denote a set of subjects $\{1, 2, \dots, S\}$, $1{:}T$ to denote a set of sessions $\{1, 2, \dots, T\}$, $1{:}N$ to denote a set of vertices $\{1, 2, \dots, N\}$, $1{:}L$ to denote a set of parcellation labels $\{1, 2, \dots, L\}$.

For each subject s at a particular session t, we computed the functional connectivity profile of each vertex (of the cortical mesh) by correlating the vertex's fMRI time course with the time courses of uniformly distributed cortical regions of interests (ROIs). For the GSP and HNU datasets, the preprocessed data were in fsaverage5 surface space. In this case, the ROIs consisted of 1175 vertices

approximately uniformly distributed across the two hemispheres (Yeo et al., 2011). For the HCP dataset, the preprocessed data is in fsLR32k surface space. In this case, the ROIs consisted of 1483 vertices spaced approximately uniformly distributed across the two hemispheres. Each vertex's connectivity profile was binarized (see Methods in main manuscript) and normalized to unit length. Let $X_n^{s,t}$ denote the binarized, normalized functional connectivity profile of subject s at vertex n during session t. Let D denote the total number of ROIs and hence the length of $X_n^{s,t}$. We denote the connectivity profiles from all sessions of all subjects at all cortical vertices as $X_{1:N}^{1:S,1:T}$.

Figure 1 (main text) illustrates the schematic of the multi-session hierarchical Bayesian model (MS-HBM). Following previous work (Yeo et al., 2011), the functional connectivity profile $X_n^{s,t}$ of subject s from a session t at vertex n is assumed to be generated from a von Mises-Fisher distribution,

$$p\left(X_n^{s,t} \middle| l_n^s = l, \mu_{1:L}^{s,t}, \kappa\right) = p\left(X_n^{s,t} \middle| \mu_l^{s,t}, \kappa\right) = z_D(\kappa)\exp\left(\kappa\langle X_n^{s,t}, \mu_l^{s,t}\rangle\right), \qquad (1)$$

where l_n^s is the parcellation label at vertex n of subject s, and $\langle\,,\rangle$ denote inner product. $\mu_l^{s,t}$ and κ are the mean direction and concentration parameter of the von Mises-Fisher distribution for network label l of subject s during session t. $\mu_{1:L}^{s,t}$ are the mean directions for networks 1 to L. We can think of $\mu_l^{s,t}$ as the mean connectivity profile of network label l normalized to unit length. If functional connectivity profile $X_n^{s,t}$ is similar to mean connectivity profile $\mu_l^{s,t}$ (i.e., $\langle X_n^{s,t}, \mu_l^{s,t}\rangle$ is big), then vertex n is more likely to be assigned to network l. The concentration parameter κ controls the variability of the functional connectivity profiles within each network. A higher κ results in a lower dispersion (i.e., lower variance), which means that vertices belonging to the same network are more likely to possess functional connectivity profiles that are close to the mean connectivity profile of the network. κ is assumed to be the same for all networks, subjects and sessions. Finally, $z_D(\kappa)$ is a normalization constant to ensure a valid probability distribution (Banerjee et al., 2005):

$$z_D(\kappa) = \frac{\kappa^{\frac{D-1}{2}-1}}{(2\pi)^{\frac{D-1}{2}} I_{\frac{D-1}{2}-1}(\kappa)}\ , \qquad (2)$$

where $I_{\frac{D-1}{2}-1}(\cdot)$ is the modified Bessel function of the first kind with order $\frac{D-1}{2} - 1$.

To model intra-subject functional connectivity variability, we assume a conjugate prior on the subject-specific and session-specific mean connectivity profiles $\mu_l^{s,t}$, which turns out to also be a von Mises-Fisher distribution:

$$p\left(\mu_l^{s,t} \middle| \mu_l^s, \sigma_l\right) = z_D(\sigma_l)\exp\left(\sigma_l \langle \mu_l^{s,t}, \mu_l^s \rangle\right), \qquad (3)$$

where μ_l^s and σ_l are the mean direction and concentration parameter of the von Mises-Fisher distribution for network label l of subject s. We can think of μ_l^s as the individual-specific functional connectivity profile of network l of subject s. The concentration parameter σ_l controls how much the session-specific mean direction $\mu_l^{s,t}$ of subject s during session t can deviate from the subject-specific mean direction μ_l^s. A higher σ_l would imply lower intra-subject functional connectivity variability across sessions. σ_l is network-specific but is assumed to be the same for all subjects.

To model inter-subject functional connectivity variability, we assume a conjugate prior on the subject-specific mean connectivity profiles μ_l^s, which is again a von Mises-Fisher distribution whose mean direction corresponded to the group-level mean direction μ^g:

$$p\left(\mu_l^s \middle| \mu_l^g, \epsilon_l\right) = z_D(\epsilon_l)\exp\left(\epsilon_l \langle \mu_l^s, \mu_l^g \rangle\right), \qquad (4)$$

where μ^g and ϵ_l are the mean direction and concentration parameter of the von Mises-Fisher distribution for network label l. We can think of μ_l^g as the group-level functional connectivity profile of network l. The concentration parameter ϵ_l controls how much the individual-specific connectivity profile μ_l^s can deviate from the group-level connectivity profile μ_l^g. A higher ϵ_l would imply lower inter-subject functional connectivity variability across subjects.

Because the functional connectivity profiles of individual subjects are generally very noisy, we impose a MRF prior on the hidden parcellation labels $l_{1:N}^s$

$$p(l_{1:N}^s) = \frac{1}{Z(\alpha,c)}\exp\left(\alpha \sum_{n=1}^N \log U(l_n^s|\Theta) - c\sum_{n=1}^N \sum_{m\in\mathcal{N}_n} V(l_n^s, l_m^s)\right), \qquad (5)$$

where $Z(\alpha, c)$ is a normalization term (partition function) to ensure $p(l_{1:N}^s)$ is a valid probability distribution. $\log U(l_n^s = l|\Theta) = \log \theta_{l,n}$ is a singleton potential encouraging certain vertices to be associated with certain labels. $V(l_n^s, l_m^s)$ is a pairwise potential (Potts model) encouraging neighboring vertices to have the same parcellation labels:

$$V(l_n^s, l_m^s) = \begin{cases} 0, & \text{if } l_n^s = l_m^s \\ 1, & \text{if } l_n^s \neq l_m^s \end{cases}, \qquad (6)$$

The parameters α and c are tunable parameters greater than zero, and control the tradeoffs between the various terms in the generative model. Assuming that $\alpha = 1$ and $c = 0$, then $\theta_{l,n}$ can be interpreted as the probability of label l occurring at vertex n of subject s.

S2. Model estimation

In this section, we describe how model parameters are estimated from a training set and a validation set (Section S2.1), and how the parameters can be used to parcellate a new subject (Section S2.2). Throughout the entire section, we assume that the number of networks $L = 17$ without loss of generality.

S2.1 Learning model parameters

Our goal is to estimate the model parameters $\{\epsilon_{1:L}, \sigma_{1:L}, \Theta_{1:N,1:L}, \mu_{1:L}^g, c, \alpha\}$ from a training set and a validation set of binarized and normalized functional connectivity profiles, which can then be utilized for estimating individual-specific parcellations in unseen data of new subjects (Section S2.2). As a reminder, $\epsilon_{1:L}$ is a group prior representing inter-subject functional connectivity variability, $\sigma_{1:L}$ is a group prior corresponding to intra-subject functional connectivity variability, $\Theta_{1:N,1:L}$ is a group prior representing inter-subject spatial variability and reflects the probability of a network occurring at particular spatial location, and $\mu_{1:L}^g$ is the group-level connectivity profile for each network. The parameters α and c tradeoff between various terms in the generative model. Because the partition function $Z(\alpha, c)$ (Eq. (5)) is NP-hard to compute, for computational efficiency, we first assume $\alpha = 1$, $c = 0$ in order to estimate $\{\epsilon_{1:L}, \sigma_{1:L}, \kappa, \mu_{1:L}^g\ \mu_{1:L}^{1:S}, \mu_{1:L}^{1:S,1:T}, \Theta_{1:N,1:L}\}$ from the training dataset. Under this scenario, $Z(\alpha, c) = 1$, and $\theta_{l,n}$ can be interpreted as the probability of label l occurring at vertex n of subject s. The tunable parameters α and c are then estimated in the validation set using a grid search.

S2.1.1 Estimating $\{\epsilon_{1:L}, \sigma_{1:L}, \kappa, \mu_{1:L}^g\ \mu_{1:L}^{1:S}, \mu_{1:L}^{1:S,1:T}, \Theta_{1:N,1:L}\}$ from training set

Given observed binarized, normalized functional connectivity profiles $X_{1:N}^{1:S,1:T}$ from the training set, we seek to estimate $\{\epsilon_{1:L}, \sigma_{1:L}, \kappa, \mu_{1:L}^g\ \mu_{1:L}^{1:S}, \mu_{1:L}^{1:S,1:T}, \Theta_{1:N,1:L}\}$ using Expectation-Maximization (EM). As previously explained, we assume $\alpha = 1, c = 0$.

Let $\Omega = \{\epsilon_{1:L}, \sigma_{1:L}, \kappa, \mu_{1:L}^{g}\ \mu_{1:L}^{1:S}, \mu_{1:L}^{1:S,1:T}, \Theta_{1:N,1:L}\}$. We consider the following maximum-a-posterior (MAP) estimation problem:

$$\underset{\Omega}{\operatorname{argmax}} \log p\left(\epsilon_{1:L}, \sigma_{1:L}, \kappa, \mu_{1:L}^{g}\ \mu_{1:L}^{1:S}, \mu_{1:L}^{1:S,1:T}, \Theta_{1:N,1:L} \middle| X_{1:N}^{1:S,1:T}\right). \qquad (7)$$

Assuming a uniform (improper) prior on $\{\Theta_{1:N,1:L}, \kappa, \sigma_{1:L}, \epsilon_{1:L}\}$, the MAP problem can be written as

$$\underset{\Omega}{\operatorname{argmax}} \log p\left(X_{1:N}^{1:S,1:T} \middle| \mu_{1:L}^{1:S,1:T}, \kappa, \Theta_{1:N,1:L}\right) p\left(\mu_{1:L}^{1:S,1:T} \middle| \sigma_{1:L}, \mu_{1:L}^{1:S}\right) p(\mu_{1:L}^{1:S} | \epsilon_{1:L}, \mu_{1:L}^{g}). \qquad (8)$$

We then introduce the parcellation labels $l_{1:N}^{s}$ for each subject s as latent variables, and use Jensen's inequality to define a lower bound $\mathcal{L}(\lambda, \Omega)$, where $\lambda = \lambda_{1:N,1:L}^{1:S}$ are the parameters of the q functions $q(l_{1:N}^{s}) = \prod_{n=1}^{N} q(l_{n}^{s} | \lambda_{n,1:L}^{s})$:

$$\begin{aligned}
&\log p\left(X_{1:N}^{1:S,1:T} \middle| \mu_{1:L}^{1:S,1:T}, \kappa, \Theta_{1:N,1:L}\right) p\left(\mu_{1:L}^{1:S,1:T} \middle| \sigma_{1:L}, \mu_{1:L}^{1:S}\right) p(\mu_{1:L}^{1:S} | \epsilon_{1:L}, \mu_{1:L}^{g}) \\
&= \sum_{s=1}^{S} \log\ p\left(X_{1:N}^{s,1:T} \middle| \mu_{1:L}^{s,1:T}, \kappa, \Theta_{1:N,1:L}\right) + \sum_{s=1}^{S} \sum_{l=1}^{L} \log p\left(\mu_{l}^{s,1:T} \middle| \sigma_{l}, \mu_{l}^{s}\right) p\left(\mu_{l}^{s} \middle| \epsilon_{l}, \mu_{l}^{g}\right) && (9) \\
&= \sum_{s=1}^{S} \log \sum_{l_{1:N}^{s}} p\left(X_{1:N}^{s,1:T}, l_{1:N}^{s} \middle| \mu_{1:L}^{s,1:T}, \kappa, \Theta_{1:N,1:L}\right) + \sum_{s=1}^{S} \sum_{l=1}^{L} \log p\left(\mu_{l}^{s,1:T} \middle| \sigma_{l}, \mu_{l}^{s}\right) p\left(\mu_{l}^{s} \middle| \epsilon_{l}, \mu_{l}^{g}\right) && (10) \\
&\geq \sum_{s=1}^{S} \sum_{l_{1:N}^{s}} q(l_{1:N}^{s}) \log \frac{p\left(X_{1:N}^{s,1:T}, l_{1:N}^{s} \middle| \mu_{1:L}^{s,1:T}, \kappa, \Theta_{1:N,1:L}\right)}{q(l_{1:N}^{s})} + \sum_{s=1}^{S} \sum_{l=1}^{L} \log p\left(\mu_{l}^{s,1:T} \middle| \sigma_{l}, \mu_{l}^{s}\right) p\left(\mu_{l}^{s} \middle| \epsilon_{l}, \mu_{l}^{g}\right) && (11) \\
&= \sum_{s=1}^{S} \sum_{t=1}^{T} \sum_{n=1}^{N} \sum_{l_{n}^{s}=1}^{L} \lambda_{n,l_{n}^{s}}^{s} \log p\left(X_{n}^{s,t} \middle| \mu_{l_{n}^{s}}^{s,t}, \kappa\right) + \sum_{s=1}^{S} \sum_{n=1}^{N} \sum_{l_{n}^{s}=1}^{L} \lambda_{n,l_{n}^{s}}^{s} \log \Theta_{n,l_{n}^{s}} \\
&\qquad - \sum_{s=1}^{S} \sum_{n=1}^{N} \sum_{l_{n}^{s}=1}^{L} \lambda_{n,l_{n}^{s}}^{s} \log \lambda_{n,l_{n}^{s}}^{s} + \sum_{s=1}^{S} \sum_{l=1}^{L} \left(\sum_{t=1}^{T} \log p(\mu^{s,t} | \sigma_{l}, \mu_{l}^{s}) + \log p(\mu_{l}^{s} | \epsilon_{l}, \mu_{l}^{g})\right) && (12) \\
&= \mathcal{L}(\lambda, \Omega), && (13)
\end{aligned}$$

where equality is achieved when $q(l_{1:N}^{s}) = \lambda_{1:N,1:L}^{s}$ are the posterior probability of the individual-specific parcellation of subject s given the parameters Ω. Therefore, instead of maximizing the original MAP problem (Eq. (7)), we instead maximize the lower bound:

$$\{\lambda^*, \Omega^*\} = \underset{\lambda,\Omega}{\operatorname{argmax}}\, \mathcal{L}(\lambda, \Omega). \qquad (14)$$

In the E-step, we fix $\Omega = \{\epsilon_{1:L}, \sigma_{1:L}, \kappa, \mu_{1:L}^{g}\, \mu_{1:L}^{1:S}, \mu_{1:L}^{1:S,1:T}, \Theta_{1:N,1:L}\}$, and estimate λ:

$$\lambda = \underset{\lambda}{\operatorname{argmax}}\, \mathcal{L}(\lambda, \Omega) \qquad (15)$$

$$= \underset{\lambda}{\operatorname{argmax}} \sum_{s=1}^{S}\sum_{t=1}^{T}\sum_{n=1}^{N}\sum_{l_n^s=1}^{L} \lambda_{n,l_n^s}^{s} \log p(X_n^{s,t} | \mu_{l_n^s}^{s,t}, \kappa) + \sum_{s=1}^{S}\sum_{n=1}^{N}\sum_{l_n^s=1}^{L} \lambda_{n,l_n^s}^{s} \log \Theta_{n,l_n^s}$$

$$- \sum_{s=1}^{S}\sum_{n=1}^{N}\sum_{l_n^s=1}^{L} \lambda_{n,l_n^s}^{s} \log \lambda_{n,l_n^s}^{s} + \sum_{s=1}^{S}\sum_{n=1}^{N} \eta_n^s \left(\sum_{l_n^s=1}^{L} \lambda_{n,l_n^s}^{s} - 1 \right), \qquad (16)$$

where η_n^s are the Lagrange multipliers enforcing the constraint $\sum_{l=1}^{L} \lambda_{n,l}^{s} = 1$. Optimizing Eq. (16), we get:

$$\log \lambda_{k,l}^{s} \propto \sum_{t=1}^{T} \log p\left(X_k^{s,t} \middle| \mu_l^{s,t}, \kappa\right) + \log \Theta_{k,l} \qquad (17)$$

$$= \sum_{t=1}^{T} \log z_D(\kappa) \exp\left(\kappa \langle X_n^{s,t}, \mu_l^{s,t} \rangle\right) + \log \Theta_{k,l} \qquad (18)$$

$$= T \log z_D(\kappa) + \sum_{t=1}^{T} \kappa \langle X_n^{s,t}, \mu_l^{s,t} \rangle + \log \Theta_{k,l} \qquad (19)$$

In the M-step, we fix λ and estimate Ω:

$$\Omega = \underset{\Omega}{\operatorname{argmax}}\, \mathcal{L}(\lambda, \Omega). \qquad (20)$$

By using the constraints that $\langle \mu_l^{s,t}, \mu_l^{s,t} \rangle = 1$, $\langle \mu_l^{s}, \mu_l^{s} \rangle = 1$, $\langle \mu_l^{g}, \mu_l^{g} \rangle = 1$, $\kappa > 0$, $\sigma_l > 0$, $\epsilon_l > 0$, and differentiating $\mathcal{L}(\lambda, \Omega)$ with respect to $\epsilon_{1:L}, \sigma_{1:L}, \kappa, \mu_{1:L}^{g}\, \mu_{1:L}^{1:S}, \mu_{1:L}^{1:S,1:T}, \Theta_{1:N,1:L}$, and setting the derivatives to zero, we get the following update equations:

$$\mu_l^{s,t} = \frac{\kappa \sum_{n=1}^{N} \lambda_{n,l}^{s} X_n^{s,t} + \sigma_l \mu_l^{s}}{\left\| \kappa \sum_{n=1}^{N} \lambda_{n,l}^{s} X_n^{s,t} + \sigma_l \mu_l^{s} \right\|} \qquad (21)$$

$$\kappa = \frac{(D-2)\Gamma^{\kappa}}{1 - \Gamma^{\kappa 2}} + \frac{(D-1)\Gamma^{\kappa}}{2(D-2)}, \Gamma^{\kappa} = \frac{\sum_{s=1}^{S}\sum_{t=1}^{T}\sum_{n=1}^{N}\sum_{l=1}^{L} \lambda_{n,l}^{s} \langle \mu_l^{s,t}, X_n^{s,t} \rangle}{T \sum_{s=1}^{S}\sum_{n=1}^{N}\sum_{l=1}^{L} \lambda_{n,l}^{s}} \qquad (22)$$

$$\mu_l^s = \frac{\sigma_l \sum_{t=1}^{T} \mu_l^{s,t} + \epsilon_l \mu_l^g}{\left\| \sigma_l \sum_{t=1}^{T} \mu_l^{s,t} + \epsilon_l \mu_l^g \right\|} \qquad (23)$$

$$\sigma_l = \frac{(D-2)\Gamma_l^{\sigma}}{1-\Gamma_l^{\sigma 2}} + \frac{(D-1)\Gamma_l^{\sigma}}{2(D-2)}, \Gamma_l^{\sigma} = \frac{1}{ST}\sum_{s=1}^{S}\sum_{t=1}^{T}\langle \mu_l^s, \mu_l^{s,t}\rangle \qquad (24)$$

$$\mu_l^g = \frac{\sum_{s=1}^{S} \epsilon_l \mu_l^s}{\left\| \sum_{s=1}^{S} \epsilon_l \mu_l^s \right\|} = \frac{\sum_{s=1}^{S} \mu_l^s}{\left\| \sum_{s=1}^{S} \mu_l^s \right\|} \qquad (25)$$

$$\epsilon_l = \frac{(D-2)\Gamma_l^{\epsilon}}{1-\Gamma_l^{\epsilon 2}} + \frac{(D-1)\Gamma_l^{\epsilon}}{2(D-2)}, \Gamma_l^{\epsilon} = \frac{1}{S}\sum_{s=1}^{S}\langle \mu_k^g, \mu_l^s\rangle \qquad (26)$$

$$\Theta_{n,l} = \frac{1}{S}\sum_{s=1}^{S} \lambda_{n,l}^s \ , \qquad (27)$$

where D is the length of $X_n^{s,t}$ (i.e., number of ROIs in each functional connectivity profile), S is the number of subjects, T is the number of sessions, and $\|\cdot\|$ corresponds to the l_2-norm. Therefore, the estimate of the functional connectivity profile $\mu_l^{s,t}$ (Eq. (21)) of network l of subject s during session t is the weighted sum of the average time course of vertices constituting network l of subject s during session t ($\sum_{n=1}^{N} \lambda_{n,l}^s X_n^{s,t}$) and the subject-specific mean direction μ_l^s, with weights κ and σ_l for each term, normalized to be unit norm. If σ_l is much greater than κ, then $\mu_l^{s,t}$ is more likely to be dominated by subject-specific mean direction μ_l^s, which means that the functional connectivity profile of network l is highly stable across sessions. Similarly, the estimate of the functional connectivity profile μ_l^s (Eq. (23)) of network l of subject s is the weighted sum of the average session-specific mean directions across all sessions for network l of subject s ($\sum_{t=1}^{T} \mu_l^{s,t}$) and the group-level mean direction μ_l^g, with weights σ_l and ϵ_l for each term, normalized to be unit norm. If ϵ_l is much greater than σ_l, then μ_l^s is more likely to be dominated by group-level mean direction μ_l^g, which means that the functional connectivity profile of network l is highly stable between subjects. Finally, the estimate of the group-level functional connectivity profile μ_l^g (Eq. (25)) of network l is the sum of the subject-specific mean directions across all subjects for network l ($\sum_{s=1}^{S} \mu_l^s$), normalized to be unit norm. The estimate of $\Theta_{n,l}$ (Eq. (27)) is the posterior probability of network l being assigned to vertex n, averaged across all the subjects.

Given the training set, the algorithm first estimates a group-level parcellation (Yeo et al., 2011), which is then used to initialize the EM algorithm. The EM algorithm iterates E-step (Eq. (19)) and M-step (Eqs. (21-27)) till convergence. We note that the update equations (Eqs. (21-27)) in the M-

step are dependent on each other. Therefore, within the M-step, the update equations (Eqs. (21-27) are iterated till convergence.

S2.1.2 Estimating tunable parameters c and α

In the previous subsection (Section S2.1.1), the training set was used to estimate $\Omega = \{\epsilon_{1:L}, \sigma_{1:L}, \kappa, \mu_{1:L}^{g}\ \mu_{1:L}^{1:S}, \mu_{1:L}^{1:S,1:T}, \Theta_{1:N,1:L}\}$, assuming $\alpha = 1, c = 0$. To tune the parameters c and α, we assume access to a validation set.

Recall that each subject in the validation set has multiple rs-fMRI sessions. We consider $c \in \{10,20,30,40,50,60\}$ and $\alpha \in \{100,150,200,250\}$. For a given pair of (c, α), and given $\{\epsilon_{1:L}, \sigma_{1:L}, \Theta_{1:N,1:L}, \mu_{1:L}^{g}\}$ estimated from the training set, we estimate for each subject in the validation set, the individual-specific parcellation based on a subset of rs-fMRI sessions (see Section S2.2 for algorithm). Resting-state homogeneity (Eq. (1) in main text) is then computed in the remaining rs-fMRI sessions of the validation subjects. The pair of (c, α) with the highest homogeneity in the unseen rs-fMRI sessions of the validation subjects is then utilized for parcellating new subjects.

In the case of the GSP data, the optimal pair of parameters is $c = 30$ and $\alpha = 200$. In the case of the HCP data, the optimal pair of parameters is $c = 40$ and $\alpha = 200$. Note that we do expect the parameters to be different between the GSP and HCP datasets because of resolution differences between the fsaverage5 and fs_LR32k surface meshes.

Throughout the paper (main text), the reported quality (Figures 5, 6 and 7) of the individual-specific parcellations was evaluated using subjects not used to tune the parameters. For example, in the case of the CoRR-HNU subjects (Figures 5 and 6), the model parameters were estimated from the GSP training and validation sets. In the case of the HCP data (Figures 6 and 7), model parameters were estimated the HCP training and validation sets, while the reported quality of the individual-specific parcellations was evaluated using the HCP test set.

S2.2 Individual-level parcellation estimation

Using parameters $\{\epsilon_{1:L}, \sigma_{1:L}, \Theta_{1:N,1:L}, \mu_{1:L}^{g}\}$ estimated from the training set (Section S2.1.1), and for a particular pair of (c, α), we can estimate the individual-specific parcellation $l_{1:N}^{s}$ of a new subject s with T sessions by employing the variational Bayes expectation maximization (VBEM) algorithm.

Let $\Psi = \{\kappa, \mu_{1:L}^{s,1:T}, \mu_{1:L}^{s}\}$. We consider the following maximum-a-posterior (MAP) estimation problem:

$$\underset{\Psi}{\operatorname{argmax}} \log p\left(\kappa, \mu_{1:L}^{s,1:T}, \mu_{1:L}^{s} \middle| X_{1:N}^{s,1:T}, \epsilon_{1:L}, \sigma_{1:L}, \mu_{1:L}^{g}, \Theta_{1:N,1:L} \right). \tag{28}$$

Assuming a uniform (improper) prior on κ, and by introducing the parcellation labels $l_{1:N}^{s}$ of the new subject s as latent variables, the lower bound $\mathcal{L}(\lambda, \Psi)$ of the MAP problem (Eq. (28)) can be written as:

$$\begin{aligned}\mathcal{L}(\lambda^s, \Psi) = &\sum_{t=1}^{T}\sum_{n=1}^{N}\sum_{l_n^s=1}^{L} \lambda_{n,l_n^s}^{s} \log p\left(X_n^{s,t} \middle| \mu_{l_n^s}^{s,t}, \kappa\right) + \alpha \sum_{n=1}^{N}\sum_{l_n^s=1}^{L} \lambda_{n,l_n^s}^{s} \log \Theta_{n,L_n^s} \\ &- c \sum_{n=1}^{N}\sum_{m\in\mathcal{N}_n}\sum_{l_n^s=1}^{L}\sum_{l_m^s=1}^{L} \lambda_{n,l_n^s}^{s} \lambda_{m,l_m^s}^{s} V(l_n^s, l_m^s) \\ &- \sum_{n=1}^{N}\sum_{l_n^s=1}^{L} \lambda_{n,l_n^s}^{s} \log \lambda_{n,l_n^s}^{s} + \sum_{l=1}^{L}\left(\sum_{t=1}^{T} \log p(\mu^{s,t} | \sigma_l, \mu_l^s) + \log p\left(\mu_l^s \middle| \epsilon_l, \mu_l^g\right)\right),\end{aligned} \tag{29}$$

where equality is achieved when λ^s is the posterior probability of the individual-specific parcellation of subject s given the parameters Ψ. Similar to Section S2.1.1, we can maximize the lower bound (Eq. (29)) by iteratively updating λ^s and Ψ. Unlike Section S2.1.1, we cannot compute the exact posterior probability λ^s because of the pairwise potentials in the Markov random field (Wainwright and Jordan, 2008). Using the mean-field approximation (Wainwright and Jordan, 2008), an approximate posterior probability λ^s is estimated in the variational E-step, while Ψ is updated in the variational M-step.

More specifically, in the variational E-step, Ψ is fixed and λ^s is estimated as follows:

$$\log \lambda_{n,l}^{s} \propto T \log z_D(\kappa) + \sum_{t=1}^{T} \kappa \langle X_n^{s,t}, \mu_l^{s,t} \rangle - 2c \sum_{m\in\mathcal{N}_n}\sum_{l_m^s=1}^{L} \lambda_{m,l_m^s}^{s} V(l_n^s, l_m^s) + \alpha \log \Theta_{n,l}. \tag{30}$$

In the variational M-step, λ^s is fixed and $\Psi = \{\kappa, \mu_{1:L}^{s,1:T}, \mu_{1:L}^{s}\}$ is estimated as follows:

$$\mu_l^{s,t} = \frac{\kappa \sum_{n=1}^{N} \lambda_{n,l}^{s} X_n^{s,t} + \sigma_l \mu_l^s}{\left\| \kappa \sum_{n=1}^{N} \lambda_{n,l}^{s} X_n^{s,t} + \sigma_l \mu_l^s \right\|} \tag{31}$$

$$\kappa = \frac{(D-2)\Gamma^{\kappa}}{1-\Gamma^{\kappa 2}} + \frac{(D-1)\Gamma^{\kappa}}{2(D-2)}, \Gamma^{\kappa} = \frac{\sum_{s=1}^{S}\sum_{t=1}^{T}\sum_{n=1}^{N}\sum_{l=1}^{L} \lambda_{n,l}^{s} \langle \mu_l^{s,t}, X_n^{s,t} \rangle}{T \sum_{s=1}^{S}\sum_{n=1}^{N}\sum_{l=1}^{L} \lambda_{n,l}^{s}} \tag{32}$$

$$\mu_l^s = \frac{\sigma_l \sum_{t=1}^{T} \mu_l^{s,t} + \epsilon_l \mu_l^g}{\left\| \sigma_l \sum_{t=1}^{T} \mu_l^{s,t} + \epsilon_l \mu_l^g \right\|} . \qquad (33)$$

Once the VBEM algorithm converges, vertex n of subject s will be assigned to label l with the highest (approximate) posterior probability.

Supplemental Results

Description	HCP field
Visual Episodic Memory	PicSeq_Unadj
Cognitive flexibility (DCCS)	CardSort_Unadj
Inhibition (Flanker task)	Flanker_Unadj
Fluid Intelligence (PMAT)	PMAT24_A_CR
Vocabulary (pronounciation)	ReadEng_Unadj
Vocabulary (picture matching)	PicVocab_Unadj
Processing Speed	ProcSpeed_Unadj
Delay Discounting	DDisc_AUC_40K
Spatial orientation	VSPLOT_TC
Sustained Attention - Sens.	SCPT_SEN
Sustained Attention - Spec.	SCPT_SPEC
Verbal Episodic Memory	IWRD_TOT
Working Memory (list sorting)	ListSort_Unadj
Cognitive status (MMSE)	MMSE_Score
Sleep quality (PSQI)	PSQI_Score
Walking endurance	Endurance_Unadj
Walking Speed	GaitSpeed_Comp
Manual dexterity	Dexterity_Unadj
Grip strength	Strength_Unadj
Odor identifcaiton	Odor_Unadj
Pain Interference Survey	PainInterf_Tscore
Taste intensity	Taste_Unadj
Contrast Sensitivity	Mars_Final
Emotional Face Matching	Emotion_Task_Face_Acc
Arithmetic	Language_Task_Math_Avg_Difficulty_Level
Story comprehension	Language_Task_Story_Avg_Difficulty_Level
Relational processing	Relational_Task_Acc
Social Cognition - random	Social_Task_Perc_Random
Social Cognition - interaction	Social_Task_Perc_TOM
Working Memory (n-back)	WM_Task_Acc
Agreeableness (NEO)	NEOFAC_A

Table S1. Descriptions and corresponding HCP field names of 58 behavioral measures utilized in Figures 8, S6 and S7.

Description	HCP field
Openness (NEO)	NEOFAC_O
Conscientiousness (NEO)	NEOFAC_C
Neuroticism (NEO)	NEOFAC_N
Extroversion (NEO)	NEOFAC_E
Emot. Recog. - Total	ER40_CR
Emot. Recog. - Angry	ER40ANG
Emot. Recog. - Fear	ER40FEAR
Emot. Recog. - Happy	ER40HAP
Emot. Recog. - Neutral	ER40NOE
Emot. Recog. - Sad	ER40SAD
Anger - Affect	AngAffect_Unadj
Anger - Hostility	AngHostil_Unadj
Anger - Aggression	AngAggr_Unadj
Fear - Affect	FearAffect_Unadj
Fear - Somatic Arousal	FearSomat_Unadj
Sadness	Sadness_Unadj
Life Satisfication	LifeSatisf_Unadj
Meaning & Purpose	MeanPurp_Unadj
Positive Affect	PosAffect_Unadj
Friendship	Friendship_Unadj
Loneliness	Loneliness_Unadj
Perceived Hostility	PercHostil_Unadj
Perceived Rejection	PercReject_Unadj
Emotional Support	EmotSupp_Unadj
Instrument Support	InstruSupp_Unadj
Perceived Stress	PercStress_Unadj
Self-Efficacy	SelfEff_Unadj

Table S1 (cont.). Descriptions and corresponding HCP field names of 58 behavioral measures utilized in Figures 8, S6 and S7.

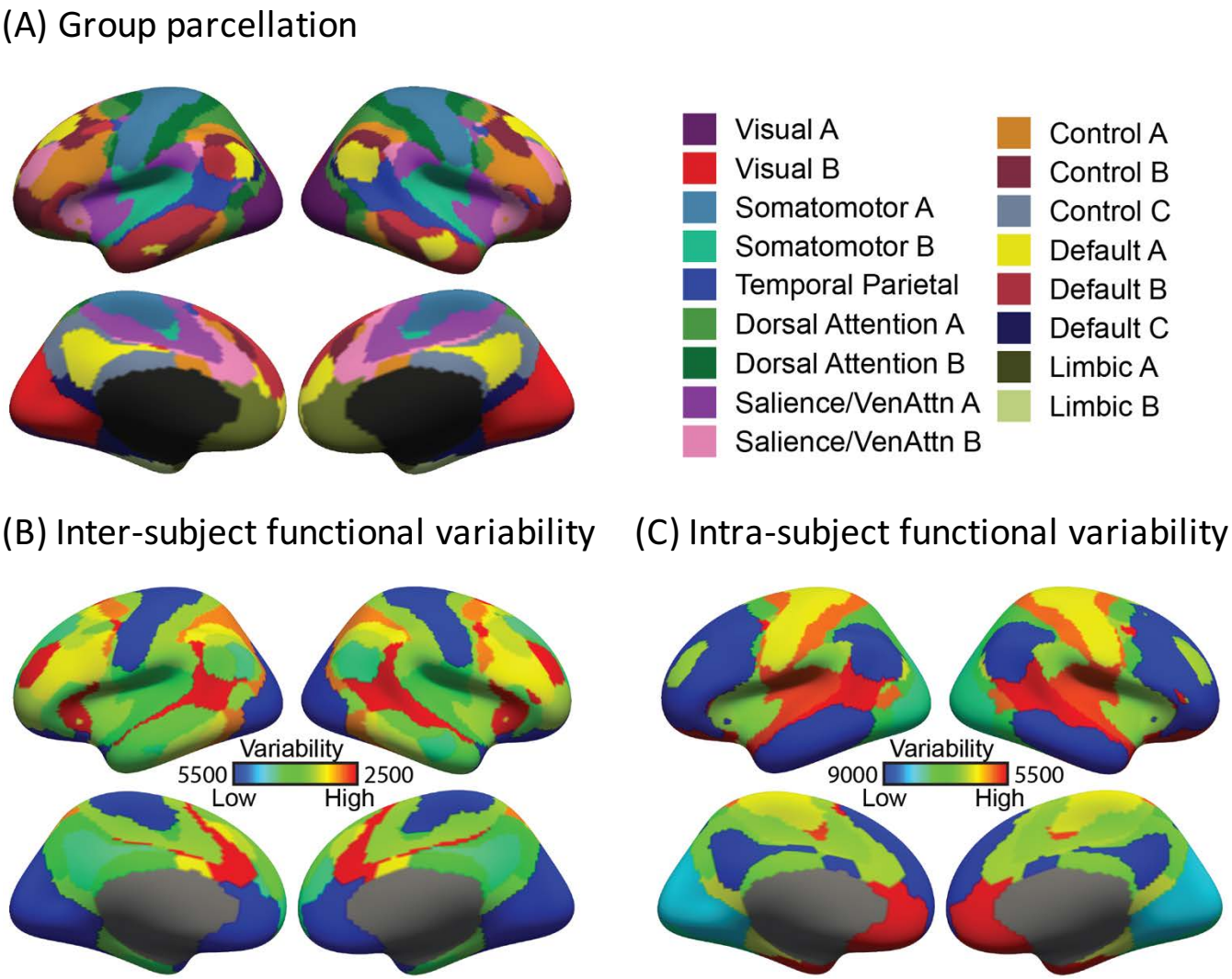

Figure S1. Sensory-motor networks exhibit lower inter-subject, but higher intra-subject, functional connectivity variability than association networks in the CoRR-HNU dataset. (A) 17-network group-level parcellation. (B) Inter-subject functional connectivity variability for different cortical networks. (C) Intra-subject functional connectivity variability for different cortical networks. Note that (B) and (C) correspond to the ϵ_l and σ_l parameters in Figure 1.

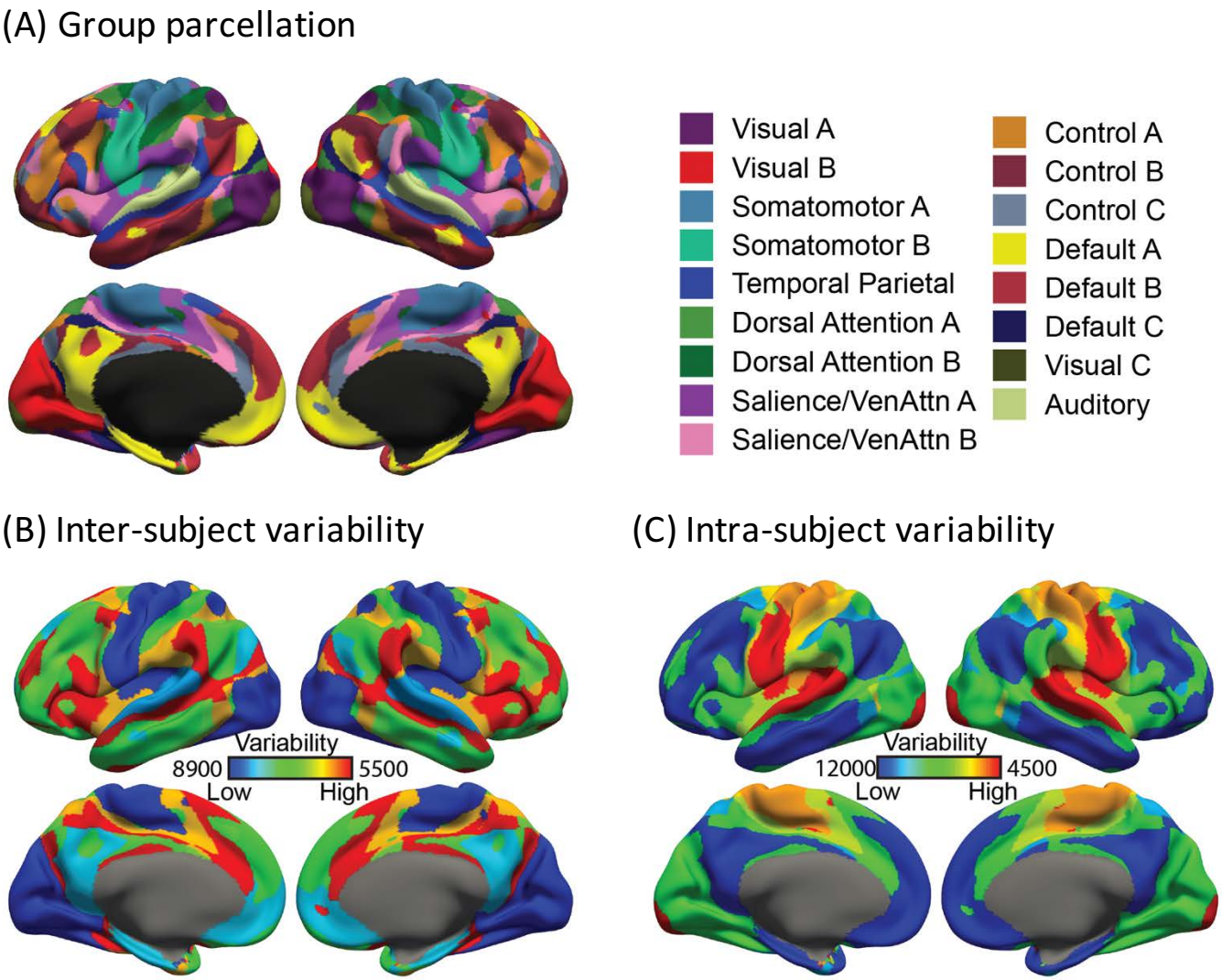

Figure S2. Sensory-motor networks exhibit lower inter-subject, but higher intra-subject, functional connectivity variability than association networks in the HCP training set. (A) 17-network group-level parcellation. (B) Inter-subject functional connectivity variability for different cortical networks. (C) Intra-subject functional connectivity variability for different cortical networks. Note that (B) and (C) correspond to the ϵ_l and σ_l parameters in Figure 1.

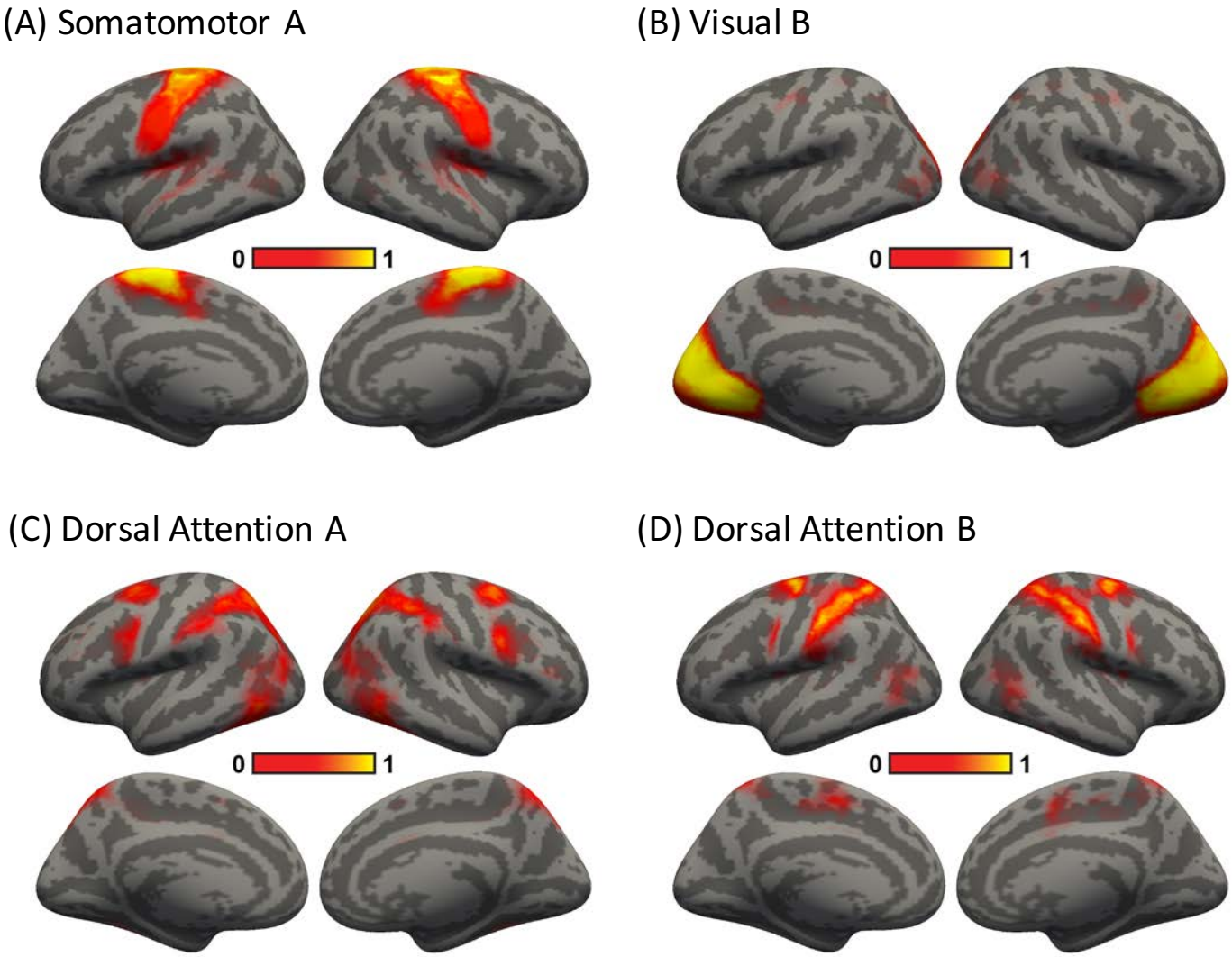

Figure S3. Sensory-motor networks are less spatially variable than association networks across subjects in the CoRR-HNU dataset. Spatial probability maps of (A) Somatomotor network A, (B) Visual network B, (C) Dorsal Attention network A, and (D) Dorsal Attention network B. A higher value (bright color) at a spatial location indicates high probability of a network appearing at that spatial location. Note that this corresponds to the Θ_l parameter in Figure 1.

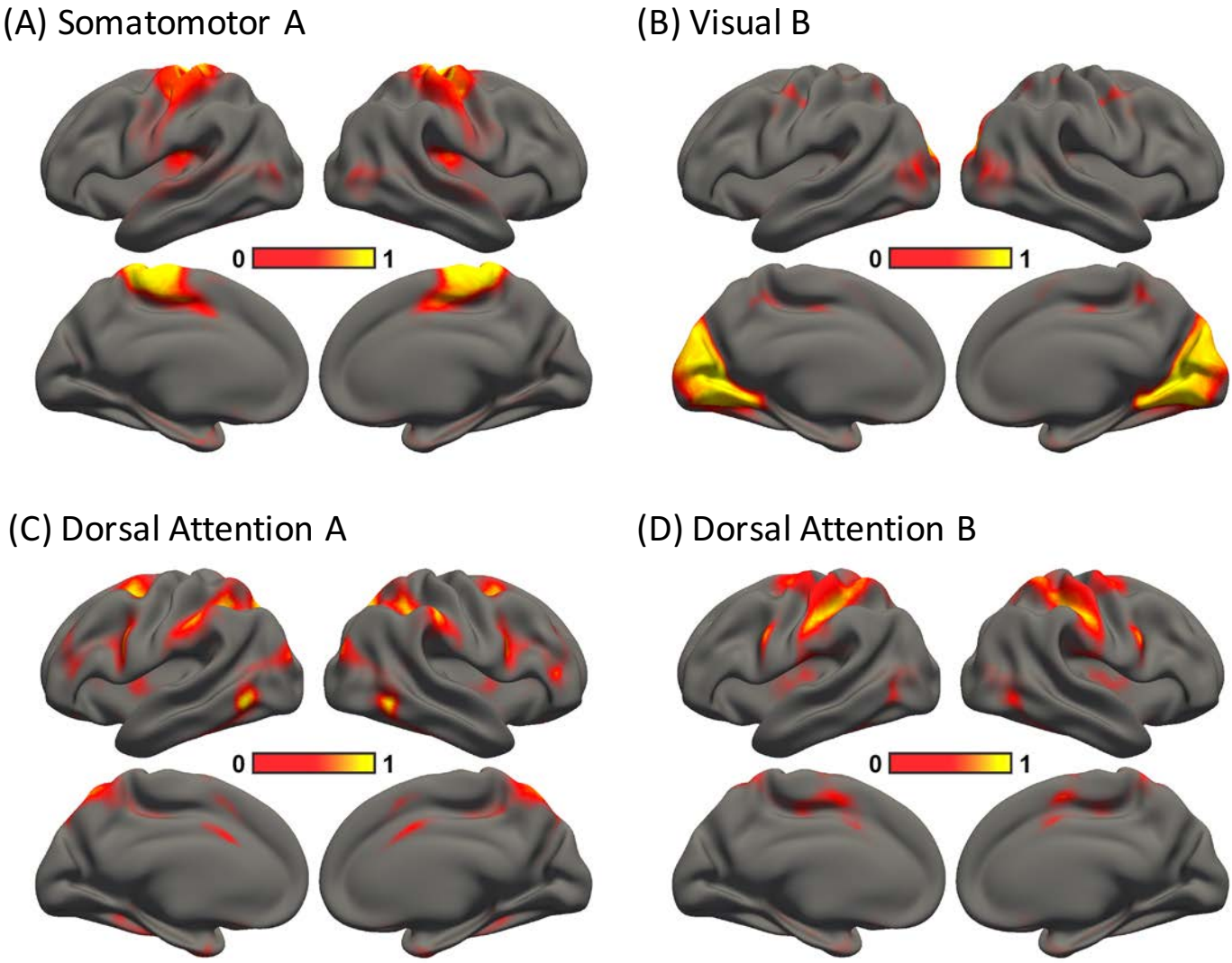

Figure S4. Sensory-motor networks are less spatially variable than association networks across subjects in the HCP training set. Spatial probability maps of (A) Somatomotor network A, (B) Visual network B, (C) Dorsal Attention network A, and (D) Dorsal Attention network B. A higher value (bright color) at a spatial location indicates high probability of a network appearing at that spatial location. Note that this corresponds to the Θ_l parameter in Figure 1.

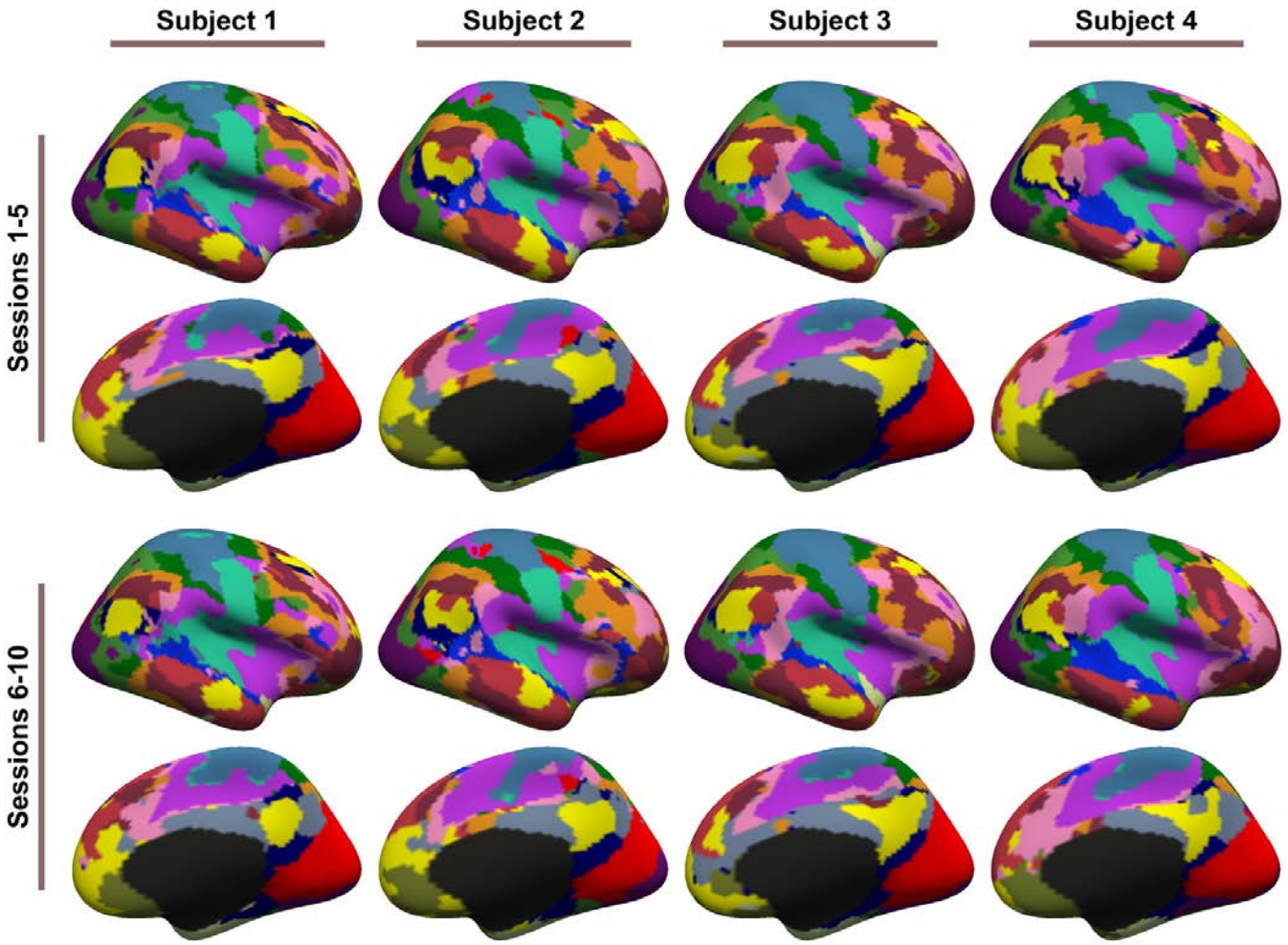

Figure S5. 17-network parcellations were estimated using sessions 1-5 and sessions 6-10 separately for each subject from the CoRR-HNU dataset. Parcellations of four representative subjects are shown here. Left hemisphere parcellations are shown in Figure 4.

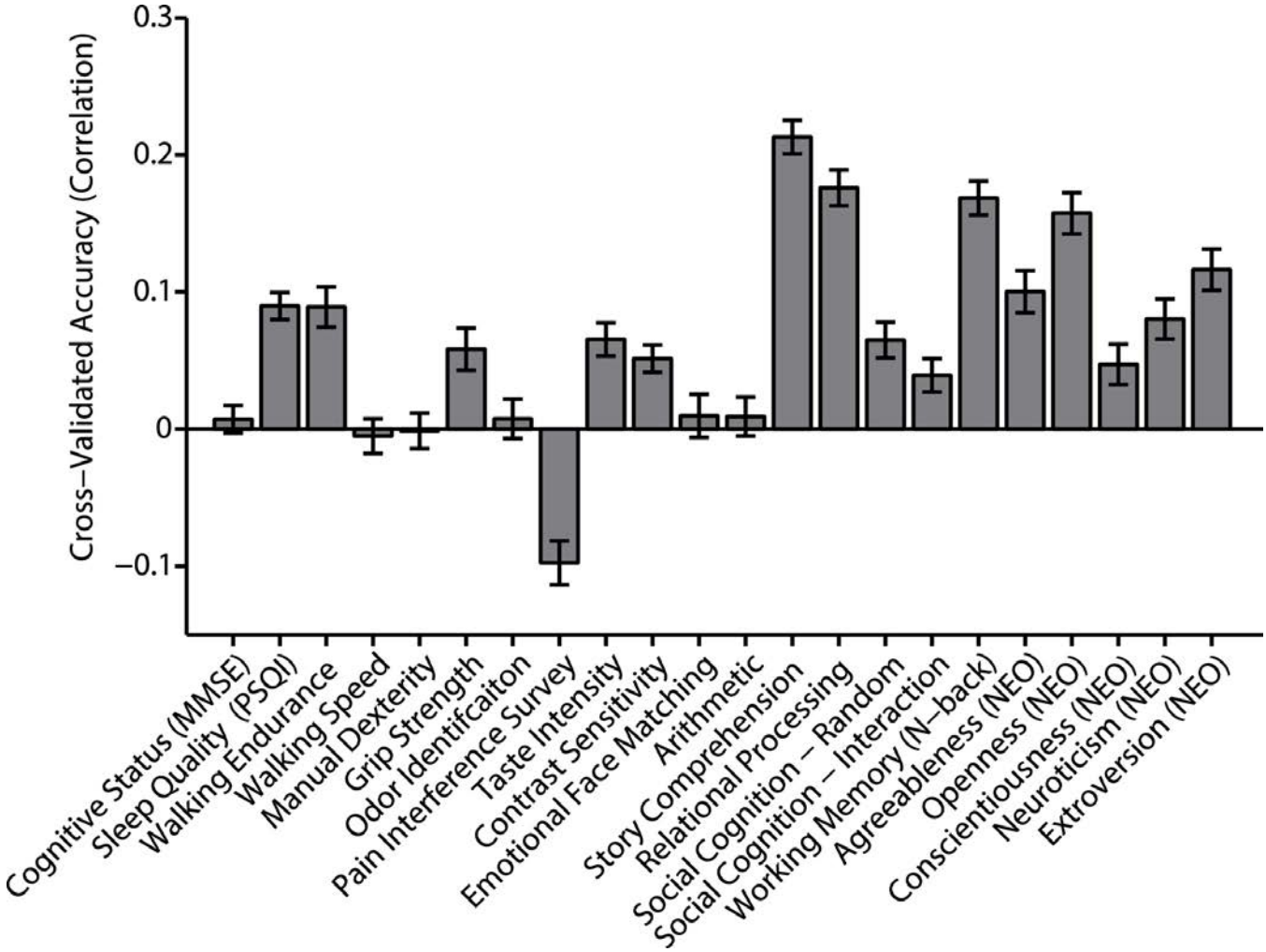

Figure S6. Prediction accuracy of 22 cognitive, emotion, personality and other non-imaging measures based on inter-subject differences in the spatial arrangement of cortical networks. Other measures are found in Figures 8 and S7.

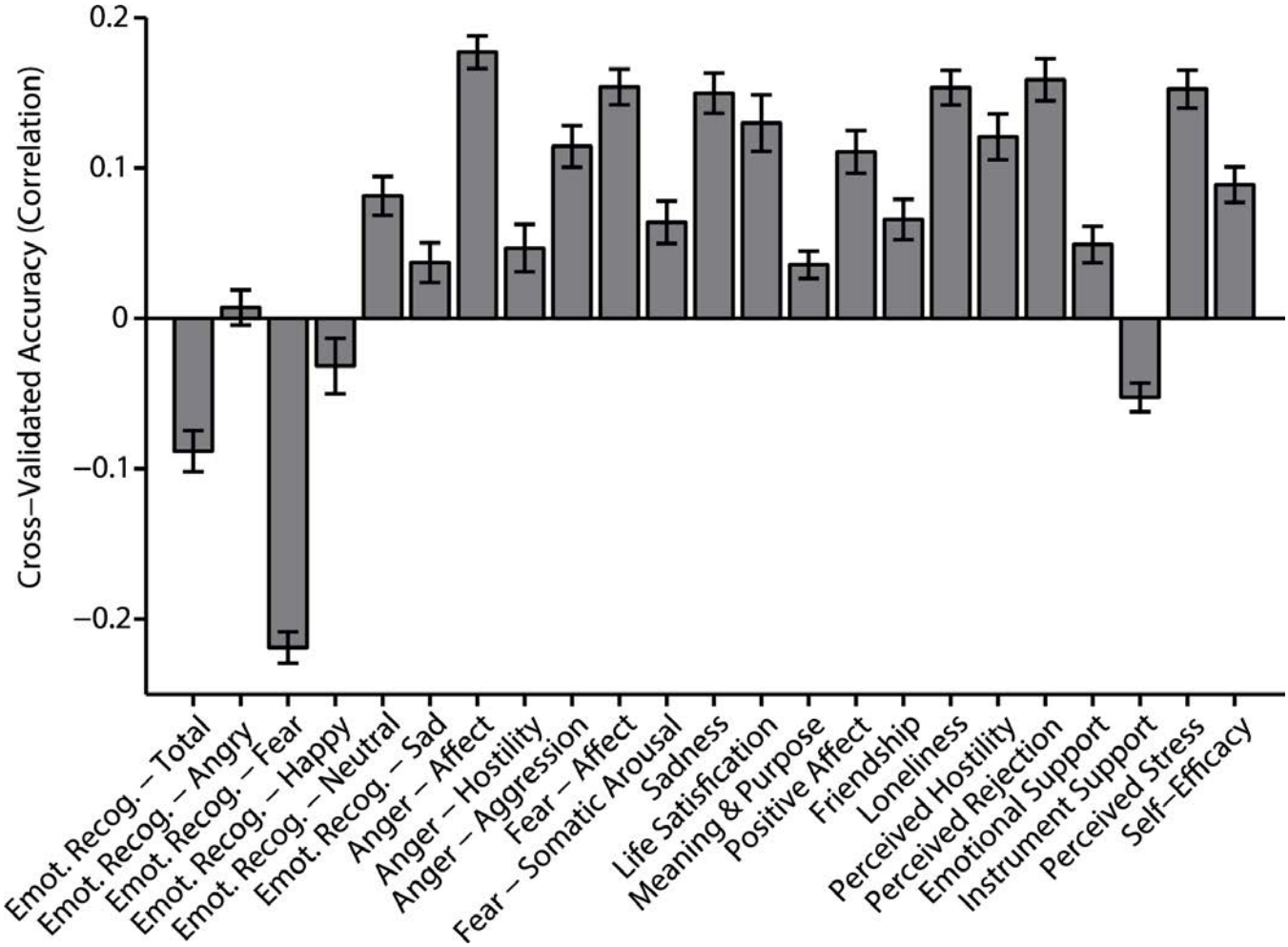

Figure S7. Prediction accuracy of 23 cognitive, emotion, personality and other non-imaging measures based on inter-subject differences in the spatial arrangement of cortical networks. Other measures are found in Figures 8 and S6. Interestingly, prediction accuracy for the emotion recognition task was poor.