Cerebral Cortex, October 2021;31: 4477–4500

https://doi.org/10.1093/cercor/bhab101
Advance Access Publication Date: 4 May 2021
Original Article

ORIGINAL ARTICLE

Individual-Specific Areal-Level Parcellations Improve Functional Connectivity Prediction of Behavior

Ru Kong[1,2,3], Qing Yang[1,2,3], Evan Gordon[4], Aihuiping Xue[1,2,3], Xiaoxuan Yan[1,2,3,5], Csaba Orban[1,2,3], Xi-Nian Zuo[6,7], Nathan Spreng[8,9,10], Tian Ge[11,12], Avram Holmes[13], Simon Eickhoff[14,15] and B.T. Thomas Yeo[1,2,3,5,12]

[1]Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117583, Singapore, [2]Centre for Sleep and Cognition (CSC) & Centre for Translational Magnetic Resonance Research (TMR), National University of Singapore, Singapore 117549, Singapore, [3]N.1 Institute for Health and Institute for Digital Medicine (WisDM), National University of Singapore, Singapore 117456, Singapore, [4]Department of Radiology, Washington University School of Medicine, St. Louis, MO 63130, USA, [5]Integrative Sciences and Engineering Programme (ISEP), National University of Singapore, Singapore 119077, Singapore, [6]State Key Laboratory of Cognitive Neuroscience and Learning/IDG McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China, [7]National Basic Public Science Data Center, Chinese Academy of Sciences, Beijing 100101, China, [8]Laboratory of Brain and Cognition, Department of Neurology and Neurosurgery, McGill University, Montreal QC H3A 2B4, Canada, [9]Departments of Psychiatry and Psychology, Neurological Institute, McGill University, Montreal QC H3A 2B4, Canada, [10]McConnell Brain Imaging Centre, Montreal Neurological Institute (MNI), McGill University, Montreal QC H3A 2B4, Canada, [11]Psychiatric & Neurodevelopmental Genetics Unit, Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA 02114, USA, [12]Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA 02129, USA, [13]Department of Psychology, Yale University, New Haven, CT 06520, USA, [14]Medical Faculty, Institute for Systems Neuroscience, Heinrich-Heine University Düsseldorf, Düsseldorf 40225, Germany and [15]Institute of Neuroscience and Medicine, Brain & Behaviour (INM-7), Research Center Jülich, Jülich 52425, Germany

Address correspondence to B.T. Thomas Yeo, N.1 & WISDM, National University of Singapore, Singapore. Email: thomas.yeo@nus.edu.sg

Abstract

Resting-state functional magnetic resonance imaging (rs-fMRI) allows estimation of individual-specific cortical parcellations. We have previously developed a multi-session hierarchical Bayesian model (MS-HBM) for estimating high-quality individual-specific network-level parcellations. Here, we extend the model to estimate individual-specific areal-level parcellations. While network-level parcellations comprise spatially distributed networks spanning the cortex, the consensus is that areal-level parcels should be spatially localized, that is, should not span multiple lobes. There is disagreement about whether areal-level parcels should be strictly contiguous or comprise multiple noncontiguous components; therefore, we considered three areal-level MS-HBM variants spanning these range of possibilities. Individual-specific MS-HBM parcellations estimated using 10 min of data generalized better than other approaches using 150 min of data to out-of-sample rs-fMRI and task-fMRI from the same individuals. Resting-state functional connectivity derived from MS-HBM parcellations also achieved the best behavioral prediction performance. Among the three MS-HBM variants, the strictly contiguous MS-HBM exhibited the best resting-state homogeneity and most uniform within-parcel task activation. In terms of behavioral prediction, the gradient-infused MS-HBM was numerically the best, but differences

among MS-HBM variants were not statistically significant. Overall, these results suggest that areal-level MS-HBMs can capture behaviorally meaningful individual-specific parcellation features beyond group-level parcellations. Multi-resolution trained models and parcellations are publicly available (https://github.com/ThomasYeoLab/CBIG/tree/master/stable_projects/brain_parcellation/Kong2022_ArealMSHBM).

Key words: behavioral prediction, brain parcellation, difference, individual, resting-state functional connectivity

Introduction

The human cerebral cortex comprises hundreds of cortical areas with distinct function, architectonics, connectivity, and topography (Kaas 1987; Felleman and Van Essen 1991; Eickhoff, Constable, et al. 2018a). These areas are thought to be organized into at least 6–20 spatially distributed large-scale networks that broadly subserve distinct aspects of human cognition (Goldman-Rakic 1988; Mesulam 1990; Smith et al. 2009; Bressler and Menon 2010; Uddin et al. 2019). Accurate parcellation of the cerebral cortex into areas and networks is therefore an important problem in systems neuroscience. The advent of in vivo noninvasive brain imaging techniques, such as functional magnetic resonance imaging (fMRI), has enabled the delineation of cortical parcels that approximate these cortical areas (Sereno et al. 1995; Van Essen and Glasser 2014; Eickhoff, Yeo, et al. 2018b).

A widely used approach for estimating network-level and areal-level cortical parcellations is resting-state functional connectivity (RSFC). RSFC reflects the synchrony of fMRI signals between brain regions, while a participant is lying at rest without performing any explicit task, that is, resting-state fMRI (rs-fMRI; Biswal et al. 1995; Fox and Raichle 2007; Buckner et al. 2013). Most RSFC studies have focused on estimating group-level parcellations obtained by averaging data across many individuals (Power et al. 2011; Yeo et al. 2011; Craddock et al. 2012; Zuo et al. 2012; Gordon et al. 2016). These group-level parcellations have provided important insights into brain network organization but fail to capture individual-specific parcellation features (Harrison et al. 2015; Laumann et al. 2015; Braga and Buckner 2017; Gordon, Laumann, Adeyemo, et al. 2017a). Furthermore, recent studies have shown that individual-specific parcellation topography is behaviorally relevant (Salehi et al. 2018; Bijsterbosch et al. 2019; Kong et al. 2019; Mwilambwe-Tshilobo et al. 2019; Seitzman et al. 2019; Li, Wang, et al. 2019b; Cui et al. 2020), motivating significant interest in estimating individual-specific parcellations.

Most individual-specific parcellations only account for inter-subject (between-subject) variability, but not intra-subject (within-subject) variability. However, inter-subject and intra-subject RSFC variability can be markedly different across brain regions (Mueller et al. 2013; Chen et al. 2015; Laumann et al. 2015). For example, the sensory-motor cortex exhibits low inter-subject variability, but high intra-subject variability (Mueller et al. 2013; Laumann et al. 2015). Therefore, it is important to consider both inter-subject and intra-subject variability when estimating individual-specific parcellations (Mejia et al. 2015, 2018; Kong et al. 2019). We have previously proposed a multi-session hierarchical Bayesian model (MS-HBM) of individual-specific network-level parcellation that accounted for both inter-subject and intra-subject variability (Kong et al. 2019). We demonstrated that compared with several alternative approaches, individual-specific MS-HBM networks generalized better to new resting-fMRI and task-fMRI data from the same individuals (Kong et al. 2019).

In this study, we extend the network-level MS-HBM to estimate individual-specific areal-level parcellations. While network-level parcellations comprise spatially distributed networks spanning the cortex, the consensus is that areal-level parcels should be spatially localized (Kaas 1987; Amunts and Zilles 2015), that is, an areal-level parcel should not span multiple cortical lobes. Consistent with invasive studies (Amunts and Zilles 2015), most areal-level parcellation approaches estimate spatially contiguous parcels (Shen et al. 2013; Honnorat et al. 2015; Gordon et al. 2016; Chong et al. 2017). However, a few studies have suggested that individual-specific areal-level parcels can be topologically disconnected (Glasser et al. 2016; Li, Wang, et al. 2019b). For example, according to Glasser et al. (2016), area 55b might comprise two disconnected, but spatially close, components in some individuals. Given the lack of consensus, we considered three different spatial localization priors. Across the three priors, the resulting parcels ranged from being strictly contiguous to being spatially localized with multiple noncontiguous components.

We compared MS-HBM areal-level parcellations with three other approaches (Laumann et al. 2015; Schaefer et al. 2018; Li, Wang, et al. 2019b) in terms of their generalizability to out-of-sample rs-fMRI and task-fMRI from the same individuals. Furthermore, a vast body of literature has shown that RSFC derived from group-level parcellations can be used to predict human behavior (Hampson et al. 2006; Finn et al. 2015; Rosenberg et al. 2016; Li, Kong, et al. 2019a). Therefore, we also investigated whether RSFC derived from individual-specific MS-HBM parcellations could improve behavioral prediction compared with two other parcellation approaches (Schaefer et al. 2018; Li, Wang, et al. 2019b).

Methods

Overview

We proposed the spatially constrained MS-HBM to estimate individual-specific areal-level parcellations. The model distinguished between inter-subject and intra-subject functional connectivity variability, while incorporating spatial contiguity constraints. Three different contiguity constraints were considered: distributed MS-HBM (dMS-HBM), contiguous MS-HBM (cMS-HBM), and gradient-infused MS-HBM (gMS-HBM). The resulting MS-HBM parcels ranged from being strictly contiguous (cMS-HBM) to being spatially localized with multiple topologically disconnected components (dMS-HBM). Subsequent analyses proceeded in four stages. First, we explored the pattern of inter-subject and intra-subject functional variability across the cortex. Second, we examined the intra-subject reproducibility and inter-subject similarity of MS-HBM parcellations on two different datasets. Third, the MS-HBM was compared with three other approaches using new rs-fMRI and task-fMRI data from the

same participants. Finally, we investigated whether functional connectivity of individual-specific parcellations could improve behavioral prediction.

Multi-session rs-fMRI Datasets

The Human Connectome Project (HCP) S1200 release (Van Essen, Ugurbil, et al. 2012a; Smith et al. 2013) comprised structural MRI, rs-fMRI, and task-fMRI of 1094 young adults. All imaging data were collected on a custom-made Siemens 3T Skyra scanner using a multiband sequence. Each participant went through two fMRI sessions on two consecutive days. Two rs-fMRI runs were collected in each session. Each fMRI run was acquired at 2 mm isotropic resolution with a time repetition (TR) of 0.72 s and lasted for 14 min and 33 s. The structural data consisted of one 0.7 mm isotropic scan for each participant.

The Midnight Scanning Club (MSC) multi-session dataset comprised structural MRI, rs-fMRI, and task-fMRI from 10 young adults (Gordon, Laumann, Gilmore, et al. 2017b; Gratton et al. 2018). All imaging data were collected on a Siemens Trio 3T MRI scanner using a 12-channel head matrix coil. Each participant was scanned for 10 sessions of rs-fMRI data. One rs-fMRI run was collected in each session. Each fMRI run was acquired at 4 mm isotropic resolution with a TR of 2.2 s and lasted for 30 min. The structural data were collected across two separate days and consisted of four 0.8 mm isotropic T1-weighted images and four 0.8 mm isotropic T2-weighted images.

It is worth noting some significant acquisition differences between the two datasets, including scanner type (e.g., Skyra vs. Trio), acquisition sequence (e.g., multiband vs. non-multiband), and scan time (e.g., day vs. midnight). These differences allowed us to test the robustness of our parcellation approach.

Preprocessing

Details of the HCP preprocessing can be found elsewhere (Van Essen, Ugurbil, et al. 2012a; Glasser et al. 2013; Smith et al. 2013; HCP S1200 manual). Of particular importance is that the rs-fMRI data have been projected to the fs_LR32k surface space (Van Essen, Glasser, et al. 2012b), smoothed by a Gaussian kernel with 2 mm full-width at half-maximum (FWHM), denoised with ICA-FIX (Griffanti et al. 2014; Salimi-Khorshidi et al. 2014), and aligned with MSMAll (Robinson et al. 2014). To eliminate global and head motion-related artifacts (Burgess et al. 2016; Siegel et al. 2017), additional nuisance regression and censoring were performed (Kong et al. 2019; Li, Kong, et al. 2019a). Nuisance regressors comprised the global signal and its temporal derivative. Runs with more than 50% censored frames were removed. Participants with all four runs remaining ($N = 835$) were considered.

In the case of the MSC dataset, we utilized the preprocessed rs-fMRI data of nine subjects on fs_LR32k surface space. Preprocessing steps included slice time correction, motion correction, field distortion correction, motion censoring, nuisance regression, and bandpass filtering (Gordon, Laumann, Gilmore, et al. 2017b). Nuisance regressors comprised whole brain, ventricular and white matter signals, as well as motion regressors derived from Volterra expansion (Friston et al. 1996). The surface data were smoothed by a Gaussian kernel with 6 mm FWHM. One participant (MSC08) exhibited excessive head motion and self-reported sleep (Gordon, Laumann, Gilmore, et al. 2017b; Seitzman et al. 2019) and was thus excluded from subsequent analyses.

Functional Connectivity Profiles

As explained in the previous section, the preprocessed rs-fMRI data from the HCP and MSC datasets have been projected onto fs_LR32K surface space, comprising 59 412 cortical vertices. A binarized connectivity profile of each cortical vertex was then computed as was done in our previous study (Kong et al. 2019). More specifically, we considered 1483 regions of interest (ROIs) consisting of single vertices uniformly distributed across the fs_LR32K surface meshes (Kong et al. 2019). For each rs-fMRI run of each participant, the Pearson's correlation between the fMRI time series at each spatial location (59 412 vertices) and the 1483 ROIs was computed. Outlier volumes were ignored when computing the correlations. The $59\,412 \times 1483$ RSFC (correlation) matrices were then binarized by keeping the top 10% of the correlations to obtain the final functional connectivity profile (Kong et al. 2019).

We note that because fMRI is spatially smooth and exhibits long-range correlations, therefore considering only 1483 ROI vertices (instead of all 59 412 vertices) would reduce computational and memory demands, without losing much information. To verify significant information has not been lost, the following analysis was performed. For each HCP participant, a $59\,412 \times 59\,412$ RSFC matrix was computed from the first rs-fMRI run. We then correlated every pair of rows of the RSFC matrix, yielding a $59\,412 \times 59\,412$ RSFC similarity matrix for each HCP participant. An entry in this RSFC similarity matrix indicates the similarity of the functional connectivity profiles of two cortical locations. The procedure was repeated but using the $59\,412 \times 1483$ RSFC matrices to compute the $59\,412 \times 59\,412$ RSFC similarity matrices. Finally, for each HCP participant, we correlated the RSFC similarity matrix (generated from 1483 vertices) and RSFC similarity matrix (generated from 59 412 vertices). The resulting correlations were high with $r = 0.9832 \pm 0.0041$ (mean $\pm$ SD) across HCP participants, suggesting that very little information was lost by only considering 1483 vertices.

Group-Level Parcellation

We have previously developed a set of high-quality population-average areal-level parcellations of the cerebral cortex (Schaefer et al. 2018), which we will refer to as "Schaefer2018." Although the Schaefer2018 parcellations are available in different spatial resolutions, we will mostly focus on the 400-region parcellation in this paper (Fig. 4A), given that previous work has suggested that there might be between 300 and 400 human cortical areas (Van Essen, Glasser, et al. 2012b). The 400-region Schaefer2018 parcellation will be used to initialize the areal-level MS-HBM for estimating individual-specific parcellations. The Schaefer2018 parcellation will also be used as a baseline in our experiments.

Areal-Level MS-HBM

The areal-level MS-HBM (Fig. 1A) is the same as the network-level MS-HBM (Kong et al. 2019) except for one crucial detail, that is, spatial localization prior Φ (Fig. 1A). Nevertheless, for completeness, we will briefly explain the other components of the MS-HBM, although further details can be found elsewhere (Kong et al. 2019).

We denote the binarized functional connectivity profile of cortical vertex n during session t of subject s as $X_n^{s,t}$. For example, the binarized functional connectivity profiles of 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 are illustrated in Fig. 1A

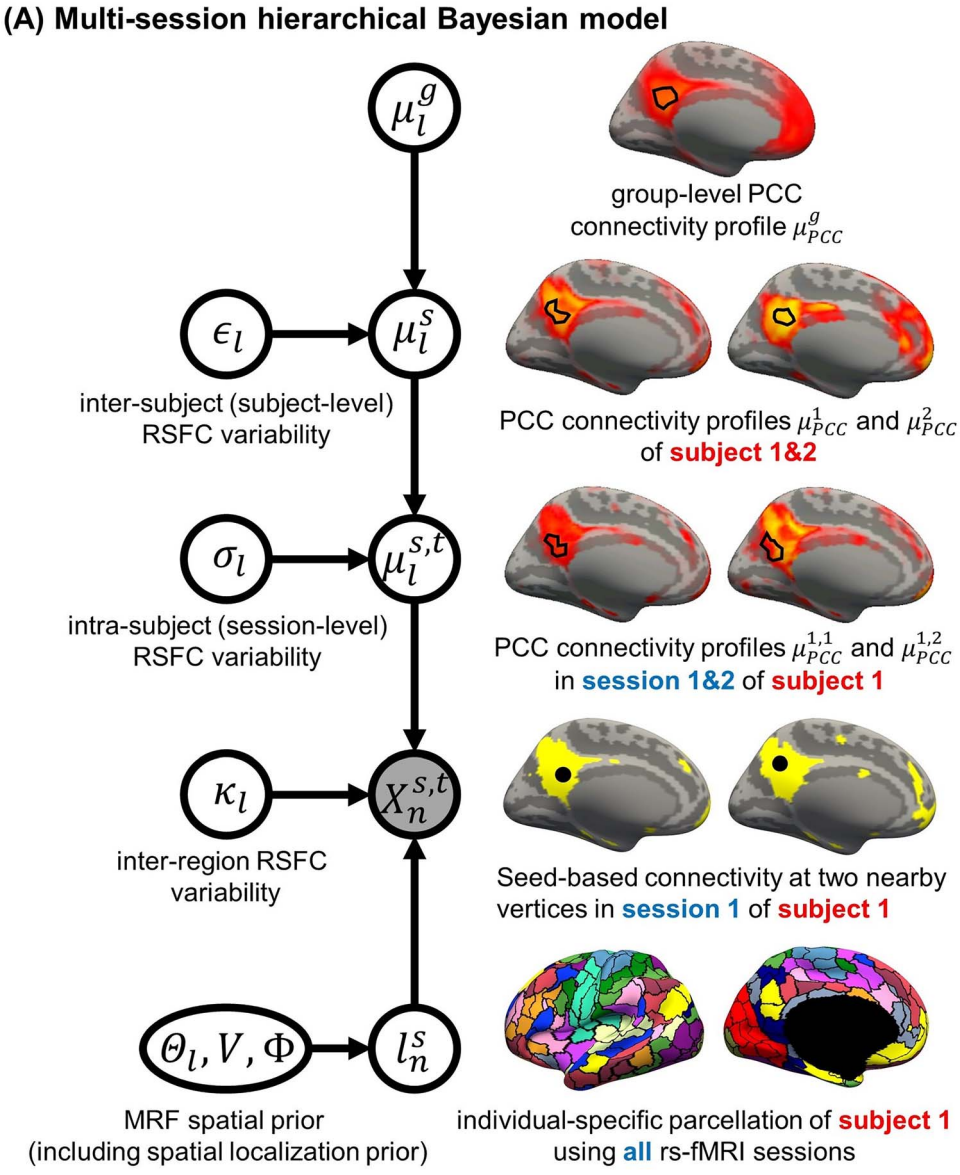

Figure 1. (*A*) MS-HBM of individual-specific areal-level parcellations. $X_n^{s,t}$ denote the RSFC profile at brain location n of subject s during rs-fMRI session t. The shaded circle indicates that $X_n^{s,t}$ are the only observed variables. The goal is to estimate the parcel label l_n^s for subject s at location n given RSFC profiles from all sessions. μ_l^g is the group-level RSFC profile of parcel l. μ_l^s is the subject-specific RSFC profile of parcel l. A large ϵ_l indicates small inter-subject RSFC variability, that is, the group-level and subject-specific RSFC profiles are very similar. $\mu_l^{s,t}$ is the subject-specific RSFC profile of parcel l during session t. A large σ_l indicates small intra-subject RSFC variability, that is, the subject-level and session-level RSFC profiles are very similar. κ_l captures inter-region RSFC variability. A large κ_l indicates small inter-region variability, that is, two locations from the same parcel exhibit very similar RSFC profiles. Finally, Θ_l captures inter-subject variability in the spatial distribution of parcels, smoothness prior V encourages parcel labels to be spatially smooth, and the spatial localization prior Φ ensures each parcel is spatially localized. The spatial localization prior Φ is the crucial difference from the original network-level MS-HBM (Kong et al. 2019). (*B*) Illustration of three different spatial localization priors. Individual-specific parcellations of the same HCP participant were estimated using dMS-HBM, cMS-HBM, and gMS-HBM. Four parcels depicted in pink, red, blue, and yellow are shown here. All four parcels estimated by dMS-HBM were spatially close together but contained two separate components. All four parcels estimated by cMS-HBM were spatially contiguous. Three parcels (pink, red, and yellow) estimated by gMS-HBM were spatially contiguous, while the blue parcel contained two separate components.

(fourth row). The shaded circle indicates that $X_n^{s,t}$ are the only observed variables. Based on the observed connectivity profiles of "all" vertices during "all" sessions of a single subject, the goal is to assign a parcel label l_n^s for each vertex n of subject s. Even though a vertex's connectivity profiles are likely to be different across fMRI sessions, the vertex's parcel label was assumed to be the same across sessions. For example, the individual-specific areal-level parcellation of the first subject using data from all available sessions is illustrated in Fig. 1A (last row).

The multiple layers of the areal-level MS-HBM explicitly differentiate inter-subject (between-subject) functional connectivity variability from intra-subject (within-subject) functional connectivity variability (ϵ_l and σ_l in Fig. 1A). The connectivity profiles of two vertices belong to the same parcel will not be identical. This variability is captured by κ_l (Fig. 1A). Some model parameters (e.g., group-level connectivity profiles) will be estimated from a training set comprising multi-session rs-fMRI data from multiple subjects. A new participant (possibly from another dataset) with single-session fMRI data could then be parcellated without access to the original training data.

The Markov random field (MRF) spatial prior (Fig. 1A last row) is important because the observed functional connectivity profiles of individual subjects are generally very noisy. Therefore, additional priors were imposed on the parcellation. First, the spatial smoothness prior V encouraged neighboring vertices (e.g., PCC and pCun) to be assigned to the same parcels. Second, the inter-subject spatial variability prior $\Theta_{l,n}$ denote the probability of parcel l occurring at a particular spatial location n. The two priors (V and $\Theta_{l,n}$) are also present in the network-level MS-HBM (Kong et al. 2019).

However, an additional spatial prior is necessary because of well-documented long-range connections spanning the cortex. Therefore, with the original MRF prior (Kong et al. 2019), brain locations with similar functional connectivity profiles could be grouped together regardless of spatial proximity. In the case of network-level MS-HBM, this is appropriate because large-scale networks are spatially distributed, for example, the default network spans frontal, parietal, temporal, and cingulate cortices. In the case of areal-level parcellations, there is the expectation that a single parcel should not span large spatial distances (Glasser et al. 2016; Gordon et al. 2016; Schaefer et al. 2018). Therefore, the areal-level MS-HBM incorporates an additional prior Φ constraining parcels to be spatially localized (Fig. 1A last row).

As mentioned in the Introduction, even though there is consensus that individual-specific areal-level parcels should be spatially localized, there are differing opinions about whether they should be spatially contiguous. Some studies have enforced spatially contiguous cortical parcels (Laumann et al. 2015; Gordon et al. 2016; Chong et al. 2017) consistent with invasive studies (Amunts and Zilles 2015). Other studies have estimated parcels that might comprise multiple spatially close components (Glasser et al. 2016; Li, Wang, et al. 2019b). For example, Glasser and colleagues suggested that area 55b might be split into two disconnected components in close spatial proximity. Given the lack of consensus, we consider three possible spatial localization priors (i.e., Φin Fig. 1A):

1. dMS-HBM. Previous studies have suggested that after registering cortical folding patterns, interindividual variability in architectonic locations are different across architectonic areas (Fischl et al. 2008). One of the most spatially variable architectonic areas is hOc5, which can be located in an adjacent sulcus away from the group-average location (Yeo et al. 2010a, 2010b). This variability corresponded to about 30 mm. Therefore, similar to Glasser et al. (2016), Φ comprises a spatial localization prior constraining each individual-specific parcel to be within 30 mm of the group-level Schaefer2018 parcel boundaries. We note that this prior only guarantees an individual-specific parcel to be spatially localized, but the parcel might comprise multiple distributed components (Fig. 1B left panel). We refer to this prior as dMS-HBM.
2. cMS-HBM. In addition to the 30 mm prior from dMS-HBM, we include a spatial localization prior encouraging vertices comprising a parcel to not be too far from the parcel center, as was done in our previous study (Schaefer et al. 2018). If this spatial contiguity prior is sufficiently strong, then all individual-specific parcels will be spatially connected (Fig. 1B middle panel). However, an overly strong prior will result overly round parcels, which is not biologically plausible (Vogt 2009). To ameliorate this issue, the estimation procedure starts with a very small weight on this spatial contiguity prior and then progressively increases the weight to ensure spatial contiguity. Thus, we refer to this prior as cMS-HBM. We note that requiring parcels to be spatially connected within an MRF framework is nontrivial; our approach is significantly less computationally expensive than competing approaches (Nowozin and Lampert 2010; Honnorat et al. 2015).
3. gMS-HBM. A well-known areal-level parcellation approach is the local gradient approach, which detects local abrupt changes (i.e., gradients) in RSFC across the cortex (Cohen et al. 2008). Our previous study (Schaefer et al. 2018) has suggested the utility of combining local gradient (Cohen et al. 2008; Gordon et al. 2016) and global clustering (Yeo et al. 2011) approaches for estimating areal-level parcellations. Therefore, we complemented the spatial contiguity prior in cMS-HBM with a prior based on local gradients in RSFC, which encouraged adjacent brain locations with gentle changes in functional connectivity to be grouped into the same parcel. In practice, we found that the gradient-infused prior, together with a very weak spatial contiguity prior, dramatically increased the number of spatially contiguous parcels (Fig. 1B right panel). Furthermore, the parcels are also less round than cMS-HBM, which is in our opinion more biologically plausible. We refer to this prior as gMS-HBM.

A more detailed mathematical explanation of the model can be found in Supplementary Methods S1. Given a dataset of subjects with multi-session rs-fMRI data, a variational Bayes expectation–maximization (VBEM) algorithm can be used to estimate the following model parameters (Kong et al. 2019): group-level parcel 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. The individual-specific areal-level parcellation of a new participant could then be generated using these estimated group-level priors without access to the original training data. Furthermore, although the model requires multi-session fMRI data for parameter estimation, it can be applied to a single-session fMRI data from a new participant (Kong et al. 2019). Details of the VBEM algorithm can be found in Supplementary Methods S2.

Characterizing Inter-subject and Intra-Subject Functional Connectivity Variability

Previous studies have shown that sensory-motor regions exhibit lower inter-subject, but higher intra-subject functional connectivity variability than association regions (Mueller et al. 2013; Laumann et al. 2015; Kong et al. 2019). Therefore, we first evaluate whether estimates of areal-level inter-subject and intra-subject variability were consistent with previous work (Fig. 2A). The HCP dataset was divided into training ($N = 40$), validation ($N = 40$), and test ($N = 755$) sets. Each HCP participant underwent two fMRI sessions on two consecutive days. Within each session, there were two rs-fMRI runs. All four runs were utilized.

The parameters of three MS-HBM variants (dMS-HBM, cMS-HBM, and gMS-HBM) were estimated. More specifically, the group-level parcel connectivity profiles μ_l^g, the inter-subject RSFC variability ϵ_l, the intra-subject RSFC variability σ_l, and inter-subject spatial variability prior Θ_l were estimated using the HCP training set (Fig. 2A). We tuned the "free" parameters (associated with the spatial smoothness prior V and spatial localization prior Φ) using the HCP validation set (Fig. 2A). The Schaefer2018 400-region group-level parcellation (Fig. 4A) was used to initialize the optimization procedure. The final trained MS-HBMs (Fig. 2A) were used in all subsequent analyses.

Intra-Subject Reproducibility and Inter-subject Similarity of MS-HBM Parcellations

Within-subject reliability is important for clinical applications (Shehzad et al. 2009; Birn et al. 2013; Zuo and Xing 2014; Zuo et al. 2019). Having verified that the spatial patterns of inter-subject and intra-subject functional connectivity variability were consistent with previous work, we further characterized the intra-subject reproducibility and inter-subject similarity of individual-specific MS-HBM parcellations (Fig. 2B). The three trained models (dMS-HBM, cMS-HBM, and gMS-HBM) were applied to the HCP test set. Individual-specific MS-HBM parcellations were independently estimated using the first two runs (day 1) and the last two runs (day 2).

To evaluate the reproducibility of individual-specific parcellations, the Dice coefficient was computed for each parcel from the two parcellations of each participant:

$$\textbf{Dice}\left(l_s^1, l_s^2\right) = \frac{2 \times \#\text{vertices that overlap between parcels } l_s^1 \text{ and } l_s^2}{\#\text{vertices in parcel } l_s^1 + \#\text{vertices in parcel } l_s^2}$$

where l_s^1 and l_s^2 are parcel l from the two parcellations of subject s. The Dice coefficient is widely used for comparing parcellation or segmentation overlap (Destrieux et al. 2010; Sabuncu et al. 2010; Birn et al. 2013; Blumensath et al. 2013; Arslan et al. 2015; Honnorat et al. 2015; Salehi et al. 2018). The Dice coefficient is equal to 1 if there is perfect overlap between parcels and zero if there is no overlap between parcels. The Dice coefficients were averaged across all participants to provide insights into regional variation in intra-subject parcel similarity. Finally, the Dice coefficients were averaged across all parcels to provide an overall measure of intra-subject parcellation reproducibility.

To evaluate inter-subject parcellation similarity, for each pair of participants, the Dice coefficient was computed for each parcel. Since there were two parcellations for each participant, there were a total of four Dice coefficients for each parcel, which were then averaged. Furthermore, the Dice coefficients were averaged across all pairs of participants to provide insights into regional variation in inter-subject parcel similarity. Finally, the dice coefficients were averaged across all parcels to provide an overall measure of inter-subject parcellation similarity.

To evaluate whether the parameters of MS-HBM algorithms from one dataset could be generalized to another dataset with different acquisition protocols and preprocessing pipelines, we used the HCP model parameters to estimate individual-specific parcellations in the MSC dataset. More specifically, the MS-HBM parcellations were independently estimated using the first five sessions and the last five sessions for each MSC participant (Fig. 2B).

Geometric Properties of MS-HBM Parcellations

The three MS-HBM variants impose different spatial priors on areal-level parcellations. To characterize the geometric properties of the MS-HBM parcels (Fig. 2C), the three trained models (dMS-HBM, cMS-HBM, and gMS-HBM) were applied to the HCP test set using all four rs-fMRI runs. We then computed two metrics to characterize the geometry of the parcellations. First, for each parcellation, the number of spatially disjoint components was computed for each parcel and averaged across all parcels. Second, for each parcellation, a roundness metric was computed for each parcel and averaged across all parcels. Here, the roundness of a parcel is defined as $1 - \frac{\#\text{parcel boundary vertices}}{\#\text{vertices contained in the parcel}}$; a larger value indicates that a parcel is rounder.

Comparison with Alternative Approaches

Here, we compared the three MS-HBM approaches (dMS-HBM, cMS-HBM, and gMS-HBM) with three alternative approaches. The first approach was to apply the Schaefer2018 400-region group-level parcellation to individual subjects. The second approach is the well-known gradient-based boundary mapping algorithm that has been widely utilized to estimate individual-specific areal-level parcellation (Laumann et al. 2015; Gordon, Laumann, Gilmore, et al. 2017b). We will refer to this second approach as "Laumann2015" (https://sites.wustl.edu/petersenschlaggarlab/resources). The third approach is the recent individual-specific areal-level parcellation algorithm of Li, Wang, et al. (2019b) (http://nmr.mgh.harvard.edu/bid/DownLoad.html), which we will refer to as "Li2019."

Evaluating the quality of individual-specific resting-state parcellations is difficult because of a lack of ground truth. Here, we considered two common evaluation metrics (Gordon et al. 2016; Chong et al. 2017; Schaefer et al. 2018; Kong et al. 2019): resting-state connectional homogeneity and task functional inhomogeneity (i.e., uniform task activation; see below). These metrics encode the principle that if an individual-specific parcellation captured the areal-level organization of the individual's cerebral cortex, then each parcel should have homogeneous connectivity and function. Furthermore, we also compared the relative utility of the different parcellation approaches for RSFC-based behavioral prediction.

Resting-State Connectional Homogeneity

Resting-state connectional homogeneity was defined as the averaged Pearson's correlations between rs-fMRI time courses of all pairs of vertices within each parcel, adjusted for parcel size and summed across parcels (Schaefer et al. 2018; Kong et al. 2019). Higher resting-state homogeneity means that

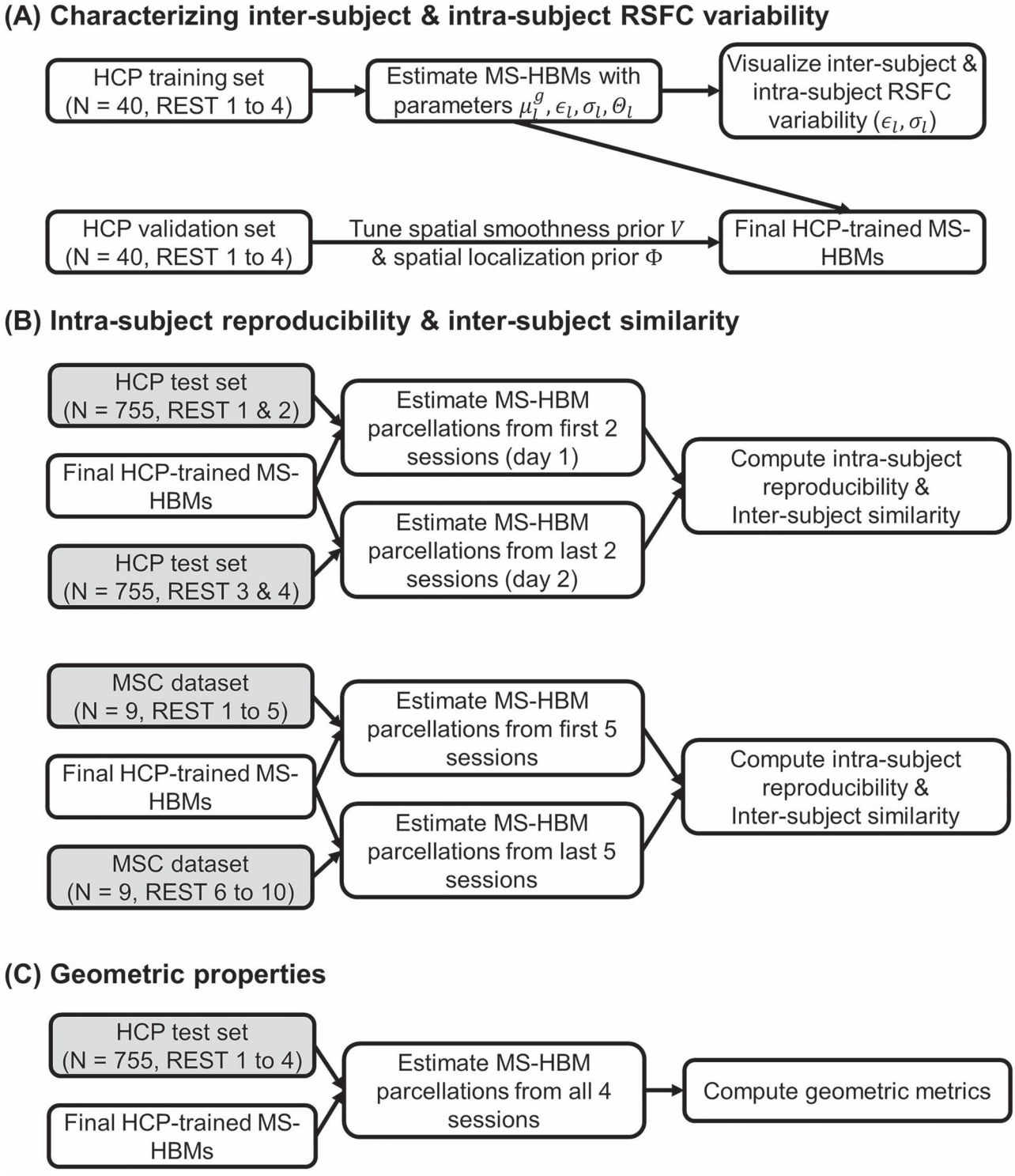

Figure 2. Flowcharts of analyses characterizing MS-HBMs. (*A*) Training MS-HBMs with HCP training and validation sets, as well as characterizing inter-subject and intra-subject RSFC variability. (*B*) Exploring intra-subject reproducibility and inter-subject similarity of MS-HBM parcellations using HCP test set and MSC dataset. (*C*) Characterizing geometric properties of MS-HBM parcellations using HCP test set. Shaded boxes (HCP test set and MSC dataset) were solely used for evaluation and not used at all for training or tuning the MS-HBM models.

vertices within the same parcel share more similar time courses. Therefore, higher resting-state homogeneity indicates better parcellation quality.

For each participant from the HCP test set ($N = 755$), we used one run to infer the individual-specific parcellation and computed resting-state homogeneity with the remaining three

runs. For the MSC dataset ($N = 9$), we used one session to infer the individual-specific parcellation and computed resting-state homogeneity with the remaining nine sessions (Fig. 3A).

Because MSC participants have large amount of rs-fMRI data (300 min), we also parcellated each MSC participant using different length of rs-fMRI data (10–150 min) and evaluated the resting-state homogeneity with the remaining five sessions. This allowed us to estimate how much the algorithms would improve with more data (Fig. 3B).

When comparing resting-state homogeneity between parcellations, the effect size (Cohen's *d*) of differences and a two-sided paired-sample *t*-test (dof = 754 for HCP, dof = 8 for MSC) were computed.

Task Functional Inhomogeneity

Task functional inhomogeneity was defined as the standard deviation (SD) of (activation) z-values within each parcel for each task contrast, adjusted for parcel size and summed across parcels (Gordon, Laumann, Gilmore, et al. 2017b; Schaefer et al. 2018). Lower task inhomogeneity means that activation within each parcel is more uniform. Therefore, lower task inhomogeneity indicates better parcellation quality. The HCP task-fMRI data consisted of seven task domains: social cognition, motor, gambling, working memory, language processing, emotional processing, and relational processing (Barch et al. 2013). The MSC task-fMRI data consisted of three task domains: motor, mixed, and memory (Gordon, Laumann, Gilmore, et al. 2017b). Each task domain contained multiple task contrasts. All available task contrasts were utilized.

For each participant from the HCP test set ($N = 755$) and MSC dataset ($N = 9$), all rs-fMRI sessions were used to infer the individual-specific parcellation (Fig. 3C). The individual-specific parcellation was then used to evaluate task inhomogeneity for each task contrast and then averaged across all available contrasts within a task domain, resulting in a single task inhomogeneity measure per task domain. When comparing between parcellations, we averaged the task inhomogeneity metric across all contrasts within a task domain before the effect size (Cohen's *d*) of differences and a two-sided paired-sample *t*-test (dof = 754 for HCP, dof = 8 for MSC) were computed for each domain.

Methodological Considerations

It is important to note that a parcellation with more parcels tends to have smaller parcel size, leading to higher resting-state homogeneity and lower task inhomogeneity. For example, if a parcel comprised two vertices, then the parcel would be highly homogeneous. In our experiments, the MS-HBM algorithms and Li2019 were initialized with the 400-region Schaefer2018 group-level parcellation, resulting in the same number of parcels as Schaefer2018, that is, 400 parcels. This allowed for a fair comparison among MS-HBMs, Li2019, and Schaefer2018.

However, parcellations estimated by Laumann2015 had a variable number of parcels across participants. Furthermore, Laumann2015 parcellations also had a significant number of vertices between parcels that were not assigned to any parcel, which has the effect of artificially increasing resting homogeneity and decreasing task inhomogeneity. Therefore, when comparing MS-HBM with Laumann2015 using resting-state homogeneity (Fig. 3A) and task inhomogeneity (Fig. 3C), we performed a post hoc processing of MS-HBM parcellations to match the number of parcels and unlabeled vertices of Laumann2015 parcellations (Supplementary Methods S3).

In addition, the Laumann2015 approach yielded different numbers of parcels within an individual with different lengths of rs-fMRI data. Therefore, Laumann2015 was also excluded from the analysis of out-of-sample resting-state homogeneity with different lengths of rs-fMRI data (Fig. 3B).

RSFC-Based Behavioral Prediction

Most studies utilized a group-level parcellation to derive RSFC for behavioral prediction (Dosenbach et al. 2010; Finn et al. 2015; Dubois et al. 2018; Weis et al. 2020; Li, Kong, et al. 2019a). Here, we investigated if RSFC derived from individual-specific parcellations can improve behavioral prediction performance. As before (He et al. 2020; Kong et al. 2019; Li, Kong, et al. 2019a), we considered 58 behavioral phenotypes measuring cognition, personality, and emotion from the HCP dataset. Three participants were excluded from further analyses because they did not have all behavioral phenotypes, resulting in a final set of 752 test participants.

The different parcellation approaches were applied to each HCP test participant using all four rs-fMRI runs (Fig. 3D). The Laumann2015 approach yielded parcellations with different numbers of parcels across participants, so there was a lack of inter-subject parcel correspondence. Therefore, we were unable to perform behavioral prediction with the Laumann2015 approach, so Laumann2015 was excluded from this analysis.

Given 400-region parcellations from different approaches (Schaefer2018; Li2019; dMS-HBM, cMS-HBM, gMS-HBM), functional connectivity was computed by correlating averaged time courses of each pair of parcels, resulting in a 400×400 RSFC matrix for each HCP test participant (Fig. 3D). Consistent with our previous work (He et al. 2020; Kong et al. 2019; Li, Wang, et al. 2019b), kernel regression was utilized to predict each behavioral measure in individual participants. Suppose y is the behavioral measure (e.g., fluid intelligence) and FC is the functional connectivity matrix of a test participant. In addition, suppose y_i is the behavioral measure (e.g., fluid intelligence) and FC_i is the individual-specific functional connectivity matrix of the ith training participant. Then kernel regression would predict the behavior of the test participant as the weighted average of the behaviors of the training participants: $y \approx \sum_{i \in \text{training set}} \text{Similarity}(FC_i, FC) y_i$. Here, $\text{Similarity}(FC_i, FC)$ is the Pearson's correlation between the functional connectivity matrices of the ith training participant and the test participant. Because the functional connectivity matrices were symmetric, only the lower triangular portions of the matrices were considered when computing the correlation. Therefore, kernel regression encodes the intuitive idea that participants with more similar RSFC patterns exhibited similar behavioral measures.

In practice, an l_2-regularization term (i.e., kernel ridge regression) was included to reduce overfitting (Supplementary Methods S4; Murphy 2012). We performed 20-fold cross-validation for each behavioral phenotype. Family structure within the HCP dataset was taken into account by ensuring participants from the same family (i.e., with either the same mother ID or father ID) were kept within the same fold and not split across folds. For each test fold, an inner-loop 20-fold cross-validation was repeatedly applied to the remaining 19 folds with different

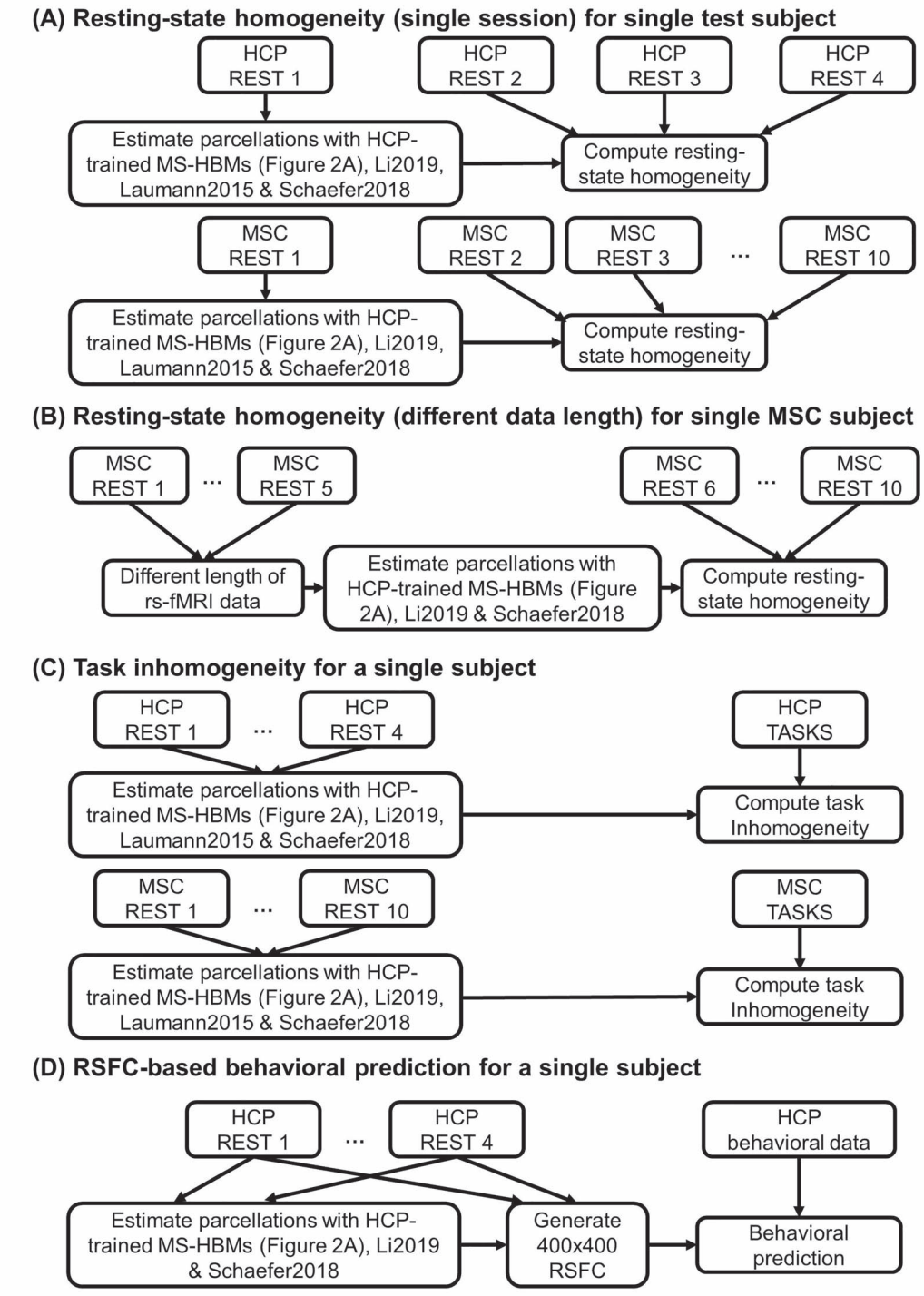

Figure 3. Flowcharts of comparisons with other algorithms. (*A*) Comparing out-of-sample resting-state homogeneity across different parcellation approaches applied to a single rs-fMRI session. (*B*) Comparing out-of-sample resting-state homogeneity across different parcellation approaches applied to different lengths of rs-fMRI data. (*C*) Comparing task inhomogeneity across different approaches. (*D*) Comparing RSFC-based behavioral prediction accuracies across different approaches. Across all analyses, MS-HBM parcellations were estimated using the trained models from Figure 2A. We remind the reader that the trained MS-HBMs were estimated using the HCP training and validation sets (Fig. 2A), which did not overlap with the HCP test set utilized in the current set of analyses. In the case of analyses (*A*) and (*B*), only a portion of rs-fMRI data was used to estimate the parcellations. The remaining rs-fMRI data were used to compute out-of-sample resting-state homogeneity. For analyses (*C*) and (*D*), all available rs-fMRI data were used to estimate the parcellations. Finally, we note that the local gradient approach (Laumann2015) does not yield a fixed number of parcels. Thus, the number of parcels is variable within an individual with different lengths of rs-fMRI data, so Laumann2015 was not considered for analysis *B*. Similarly, the number of parcels is different across participants, so the sizes of the RSFC matrices are different across participants. Therefore, Laumann2015 was also not utilized for analysis *D*.

regularization parameters. The optimal regularization parameter from the inner-loop cross-validation was then used to predict the behavioral phenotype in the test fold. Accuracy was measured by correlating the predicted and actual behavioral measure across all participants within the test fold (Finn et al. 2015; Kong et al. 2019; Li, Wang, et al. 2019b). By repeating the procedure for each test fold, each behavior yielded 20 correlation accuracies, which were then averaged across the 20 folds. Because a single 20-fold cross-validation might be sensitive to the particular split of the data into folds (Varoquaux et al. 2017), the above 20-fold cross-validation was repeated 100 times. The mean accuracy and SD across the 100 cross-validations will be reported. When comparing between parcellations, a corrected resampled *t*-test for repeated k-fold cross-validation was performed (Bouckaert and Frank 2004). We also repeated the analyses using coefficient of determination (COD) as a metric of prediction performance.

As certain behavioral measures are known to correlate with motion (Siegel et al. 2017), we regressed out age, sex, framewise displacement, DVARS, body mass index, and total brain volume from the behavioral data before kernel ridge regression. To prevent any information leak from the training data to test data, the nuisance regression coefficients were estimated from the training folds and applied to the test fold.

Code and Data Availability

Code for this work is freely available at the GitHub repository maintained by the Computational Brain Imaging Group (https://github.com/ThomasYeoLab/CBIG). The Schaefer2018 group-level parcellation and code are available here (https://github.com/ThomasYeoLab/CBIG/tree/master/stable_projects/brain_parcellation/Schaefer2018_LocalGlobal), while the areal-level MS-HBM parcellation code is available here (https://github.com/ThomasYeoLab/CBIG/tree/master/stable_projects/brain_parcellation/Kong2022_ArealMSHBM). We have also provided trained MS-HBM parameters at different spatial resolutions, ranging from 100 to 1000 parcels.

We note that the computational bottleneck for gMS-HBM is the computation of the local gradients (Laumann et al. 2015). We implemented a faster and less memory-intensive version of the local gradient computation by subsampling the functional connectivity matrices (Supplementary Methods S1.3). Computing the gradient map of a single HCP run requires 15 min and 3 GB of RAM, compared with 4 h and 40 GB of RAM in the original version. The resulting gradient maps were highly similar to the original gradient maps ($r = 0.97$). The faster gradient code can be found here (https://github.com/ThomasYeoLab/CBIG/tree/master/utilities/matlab/speedup_gradients).

The individual-specific parcellations for the HCP and MSC, together with the associated RSFC matrices, are available here (https://balsa.wustl.edu/study/show/Pr8jD and https://github.com/ThomasYeoLab/CBIG/tree/master/stable_projects/brain_parcellation/Kong2022_ArealMSHBM).

Results

Overview

Three variations of the MS-HBM with different contiguity constraints (Fig. 1) were applied to two multi-session rs-fMRI datasets to ensure that the approaches were generalizable across datasets with significant acquisition and processing differences. After confirming previous literature (Mueller et al. 2013; Laumann et al. 2015; Kong et al. 2019) that inter-subject and intra-subject RSFC variabilities were different across the cortex, we then established that the MS-HBM algorithms produced individual-specific areal-level parcellations with better quality than other approaches. Finally, we investigated whether RSFC derived from MS-HBM parcellations could be used to improve behavioral prediction.

Sensory-Motor Cortex Exhibits Lower Inter-Subject but Higher Intra-Subject Functional Connectivity Variability Than Association Cortex

The parameters of gMS-HBM, dMS-HBM and cMS-HBM were estimated using the HCP training set. Supplementary Figure S1 shows the inter-subject RSFC variability (ϵ_l) and intra-subject RSFC variability (σ_l) overlaid on corresponding Schaefer2018 group-level parcels. The pattern of inter-subject and intra-subject RSFC variability was consistent with previous work (Mueller et al. 2013; Laumann et al. 2015; Kong et al. 2019). More specifically, sensory-motor parcels exhibited lower inter-subject RSFC variability than association cortical parcels. On the other hand, association cortical parcels exhibited lower intra-subject RSFC variability than sensory-motor parcels.

Individual-Specific MS-HBM Parcellations Exhibit High Intra-Subject Reproducibility and Low Inter-Subject Similarity

To assess intra-subject reproducibility and inter-subject similarity, the three MS-HBM variants were tuned on the HCP training and validation sets and then applied to the HCP test set. Individual-specific parcellations were generated by using rs-fMRI data from day 1 (first 2 runs) and day 2 (last 2 runs) separately for each participant. All 400 parcels were present in 99% of the participants.

Figure 4 shows the inter-subject and intra-subject spatial similarity (Dice coefficient) of parcels from the three MS-HBM variants in the HCP test set. Intra-Subject reproducibility was greater than inter-subject similarity across all parcels. Consistent with our previous work on individual-specific cortical networks (Kong et al. 2019), sensory-motor parcels were more spatially similar across participants than association cortical parcels. Sensory-motor parcels also exhibited greater within-subject reproducibility than association cortical parcels.

Overall, gMS-HBM, dMS-HBM, and cMS-HBM achieved intra-subject reproducibility of 81.0%, 80.4%, and 76.1%, respectively, and inter-subject similarity of 68.2%, 68.1%, and 63.9%, respectively. We note that these metrics cannot be easily used to judge the quality of the parcellations. For example, gMS-HBM has higher intra-subject reproducibility and higher inter-subject similarity than cMS-HBM, so we cannot simply conclude that one is better than the other.

Figure 5A and Supplementary Figure S2 show the gMS-HBM parcellations of four representative HCP participants. Supplementary Figures S3 and S4 show the dMS-HBM and cMS-HBM parcellations of the same HCP participants. Consistent with previous studies of individual-specific parcellations (Glasser et al. 2016; Chong et al. 2017; Gordon, Laumann, Gilmore, et al. 2017b; Salehi et al. 2018; Seitzman et al. 2019; Li, Wang, et al. 2019b), parcel shape, size, location, and topology were variable across participants. Parcellations were highly similar within each participant with individual-specific parcel features highly preserved

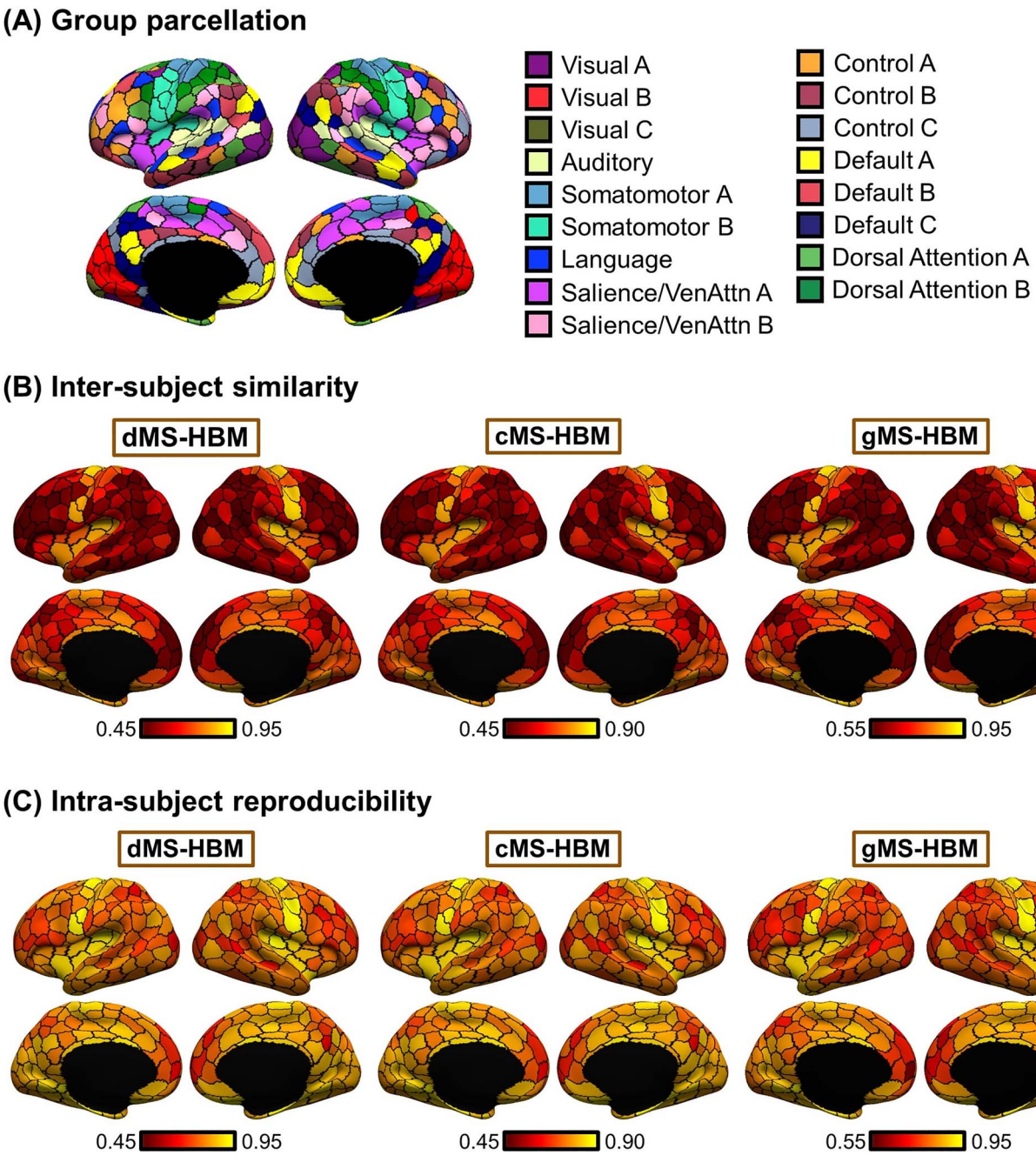

Figure 4. Individual-specific MS-HBM parcellations show high within-subject reproducibility and low across-subject similarity in the HCP test set. (*A*) The 400-region Schaefer2018 group-level parcellation. (*B*) Inter-Subject spatial similarity for different parcels. (*C*) Intra-Subject reproducibility for different parcels. Yellow color indicates higher overlap. Red color indicates lower overlap. Individual-specific MS-HBM parcellations were generated by using day 1 (first two runs) and day 2 (last two runs) separately for each participant. Sensory-motor parcels exhibited higher intra-subject reproducibility and inter-subject similarity than association parcels.

across sessions (Fig. 5B). Similar results were obtained with dMS-HBM and cMS-HBM (Fig. 5B).

The trained MS-HBM from the HCP dataset was also applied to the MSC dataset. The MS-HBM parcellations of four representative MSC participants are shown in Supplementary Figures S5–S7. Similar to the HCP dataset, the parcellations also captured unique features that were replicable across the first five sessions and the last five sessions. Overall, gMS-HBM, dMS-HBM, and cMS-HBM achieved intra-subject reproducibility of 75.5%, 73.9%, and 67.8%, respectively, and inter-subject similarity of 50.6%, 47.1%, and 42.9%, respectively.

Geometric Properties of MS-HBM Parcellations

In the HCP test set, the average number of spatially disconnected components per parcel was 1.95 ± 0.29 (mean $\pm$ SD), 1 ± 0, and 1.06 ± 0.07 for dMS-HBM, cMS-HBM, and gMS-HBM, respectively. In the case of dMS-HBM, the maximum number of spatially

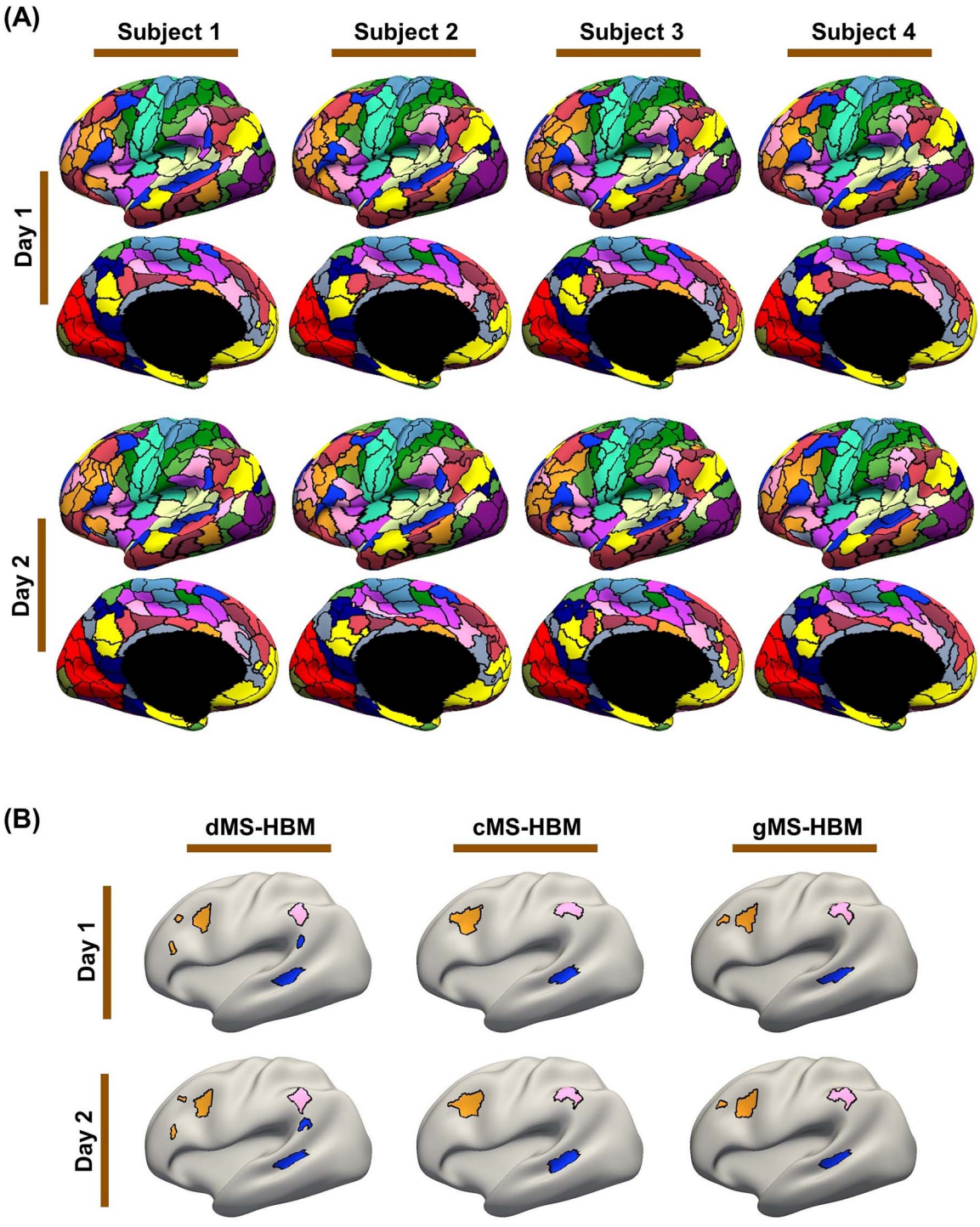

Figure 5. MS-HBM parcellations exhibit individual-specific features that are replicable across sessions. (*A*) The 400-region individual-specific gMS-HBM parcellations were estimated using rs-fMRI data from day 1 and day 2 separately for each HCP test participant. Right hemisphere parcellations are shown in Supplementary Figure S2. See Supplementary Figures S3 and S4 for dMS-HBM and cMS-HBM. (*B*) Replicable individual-specific parcellation features in a single HCP test participant for dMS-HBM, cMS-HBM, and gMS-HBM.

disconnected components (across all participants and parcels) was 11 (Supplementary Fig. S8). In the case of gMS-HBM, the maximum number of spatially disconnected components (across all participants and parcels) was 3 (Supplementary Fig. S8). On the other hand, the average roundness of the parcellations was 0.56 ± 0.02 (mean $\pm$ SD), 0.60 ± 0.01, and 0.58 ± 0.02 for dMS-HBM, cMS-HBM, and gMS-HBM, respectively. Overall, gMS-HBM parcels have much fewer spatially disconnected

components than dMS-HBM, while achieving intermediate roundness between dMS-HBM and cMS-HBM.

Individual-Specific MS-HBM Parcels Exhibit Higher Resting-State Homogeneity Than Other Approaches

Individual-specific areal-level parcellations were estimated using a single rs-fMRI session for each HCP test participant and each MSC participant. Resting-state homogeneity was evaluated using leave-out sessions in the HCP (Fig. 6A,B) and MSC (Fig. 6C,D and Supplementary Fig. S9) datasets. We note that comparisons with Laumann2015 are shown on separate plots (Fig. 6B,D) because Laumann2015 yielded different number of parcels across participants. Therefore, we matched the number of MS-HBM parcels to Laumann2015 for each participant for fair comparison (see Methods).

Across both HCP and MSC datasets, the MS-HBM algorithms achieved better homogeneity than the group-level parcellation (Schaefer2018) and two individual-specific areal-level parcellation approaches (Laumann2015 and Li2019). Compared with Schaefer2019, the three MS-HBM variants achieved an improvement ranging from 3.4% to 7.5% across the two datasets (average improvement = 5.2%, average Cohen's $d = 3.8$, largest $P = 1.9e{-}6$). Compared with Li2019, the three MS-HBM variants achieved an improvement ranging from 2.2% to 4.9% across the two datasets (average improvement = 3.4%, average Cohen's $d = 3.9$, largest $P = 5.5e{-}6$). Compared with Laumann2015, the three MS-HBM variants achieved an improvement ranging from 6.3% to 7.8% across the two datasets (average improvement = 6.7%, average Cohen's $d = 7.5$, largest $P = 1.2e{-}9$). All reported P values were significant after correcting for multiple comparisons with false discovery rate (FDR) $q < 0.05$.

Among the three MS-HBM variants, cMS-HBM achieved the highest homogeneity, while dMS-HBM was the least homogeneous. In the HCP dataset, cMS-HBM achieved an improvement of 0.19% (Cohen's $d = 0.5$, $P = 3.5e{-}38$) over gMS-HBM, and gMS-HBM achieved an improvement of 0.76% (Cohen's $d = 2.5$, $P = 3.5e{-}38$) over dMS-HBM. In the MSC dataset, cMS-HBM achieved an improvement of 1.1% (Cohen's $d = 3.5$, $P = 6.3e{-}6$) over gMS-HBM, and gMS-HBM achieved an improvement of 0.7% (Cohen's $d = 2.6$, $P = 6.1e{-}5$) over dMS-HBM. All reported P values were significant after correcting for multiple comparisons with FDR $q < 0.05$.

Individual-specific parcellations were estimated with increasing length of rs-fMRI data in the MSC dataset. Resting-state homogeneity was evaluated using leave-out sessions (Fig. 7A and Supplementary Fig. S10). We note that Laumann2015 parcellations had different number of parcels with different length of rs-fMRI data. Therefore, the resting-state homogeneity of Laumann2015 parcellations was not comparable across different length of rs-fMRI data, so the results were not shown. Because Schaefer2018 is a group-level parcellation, the parcellation stays the same regardless of the amount of data. Therefore, the performance of the Schaefer2018 group-level parcellation remained constant regardless of the amount of data. Surprisingly, the performance of the Li2019 individual-specific parcellation approach also remained almost constant regardless of the amount of data. One possible reason is that Li2019 constrained individual-specific parcels to overlap with group-level parcels. This might be an overly strong constraint, which could not be overcome with more rs-fMRI data. By contrast, the MS-HBM algorithms (dMS-HBM, cMS-HBM, and gMS-HBM) exhibited higher homogeneity with increased length of rs-fMRI data, suggesting that MS-HBM models were able to improve with more rs-fMRI data.

Furthermore, using just 10 min of rs-fMRI data, the MS-HBM algorithms achieved better homogeneity than Laumann2015 and Li2019 using 150 min of rs-fMRI data (Fig. 7B and Supplementary Fig. S10). More specifically, compared with Laumann2015 using 150 min of rs-fMRI data, dMS-HBM, cMS-HBM, and gMS-HBM using 10 min of rs-fMRI data achieved an improvement of 2.6% (Cohen's $d = 2.7$, $P = 3.6e{-}5$), 6.2% (Cohen's $d = 5.5$, $P = 1.9e{-}7$), and 5.6% (Cohen's $d = 6.1$, $P = 2.3e{-}7$), respectively. Compared with Li2019 using 150 min of rs-fMRI data, dMS-HBM, cMS-HBM, and gMS-HBM using 10 min of rs-fMRI data achieved an improvement of 0.4% (Cohen's $d = 0.4$, not significant), 2.4% (Cohen's $d = 1.9$, $P = 4.3e{-}4$), and 1.5% (Cohen's $d = 1.7$, $P = 1.0e{-}3$), respectively. All reported P values were significant after correcting for multiple comparisons with FDR $q < 0.05$.

Individual-Specific MS-HBM Parcels Exhibit Lower Task Inhomogeneity Than Other Approaches

Individual-specific parcellations were estimated using all rs-fMRI sessions from the HCP test set and MSC dataset. Task inhomogeneity was evaluated using task fMRI. Figure 8 and Supplementary Figure S11 show the task inhomogeneity of all approaches for all task domains in the MSC and HCP datasets, respectively. Compared with Schaefer2019, the three MS-HBM variants achieved an improvement ranging from 0.9% to 5.9% across all task domains and datasets (average improvement = 3.2%, average Cohen's $d = 2.4$, largest $P = 2.0e{-}3$). Compared with Li2019, the three MS-HBM variants achieved an improvement ranging from 0.8% to 5.0% across all task domains and datasets (average improvement = 2.7%, average Cohen's $d = 2.2$, largest $P = 1.8e{-}3$). Compared with Laumann2015, the three MS-HBM variants achieved an improvement ranging from 1.9% to 28.1% across all task domains and datasets (average improvement = 6.7%, average Cohen's $d = 2.3$, largest $P = 0.017$). All reported P values were significant after correcting for multiple comparisons with FDR $q < 0.05$. In the case of MSC, these improvements were observed in almost every single participant across all tasks (Fig. 8).

Among the three MS-HBM variants, cMS-HBM achieved the best task inhomogeneity, while dMS-HBM achieved the worst task inhomogeneity. Compared with gMS-HBM, cMS-HBM achieved an improvement ranging from 0.03% to 0.92% across all task domains and datasets (average improvement = 0.3%, average Cohen's $d = 0.6$, largest $P = 0.013$). Compared with dMS-HBM, gMS-HBM achieved an improvement ranging from 0.06% to 1.1% across all task domains and datasets (average improvement = 0.5%, average Cohen's $d = 1.1$, largest $P = 1.2e{-}3$). All reported P values were significant after correcting for multiple comparisons with FDR $q < 0.05$.

Functional Connectivity of Individual-Specific MS-HBM Parcels Improves Behavioral Prediction

Individual-specific parcellations were estimated using all rs-fMRI sessions from the HCP test set. The RSFC of the individual-specific parcellations was used for predicting 58 behavioral measures. We note that the number of parcels was different across participants for Laumann2015, so Laumann2015 could not be included for this analysis.

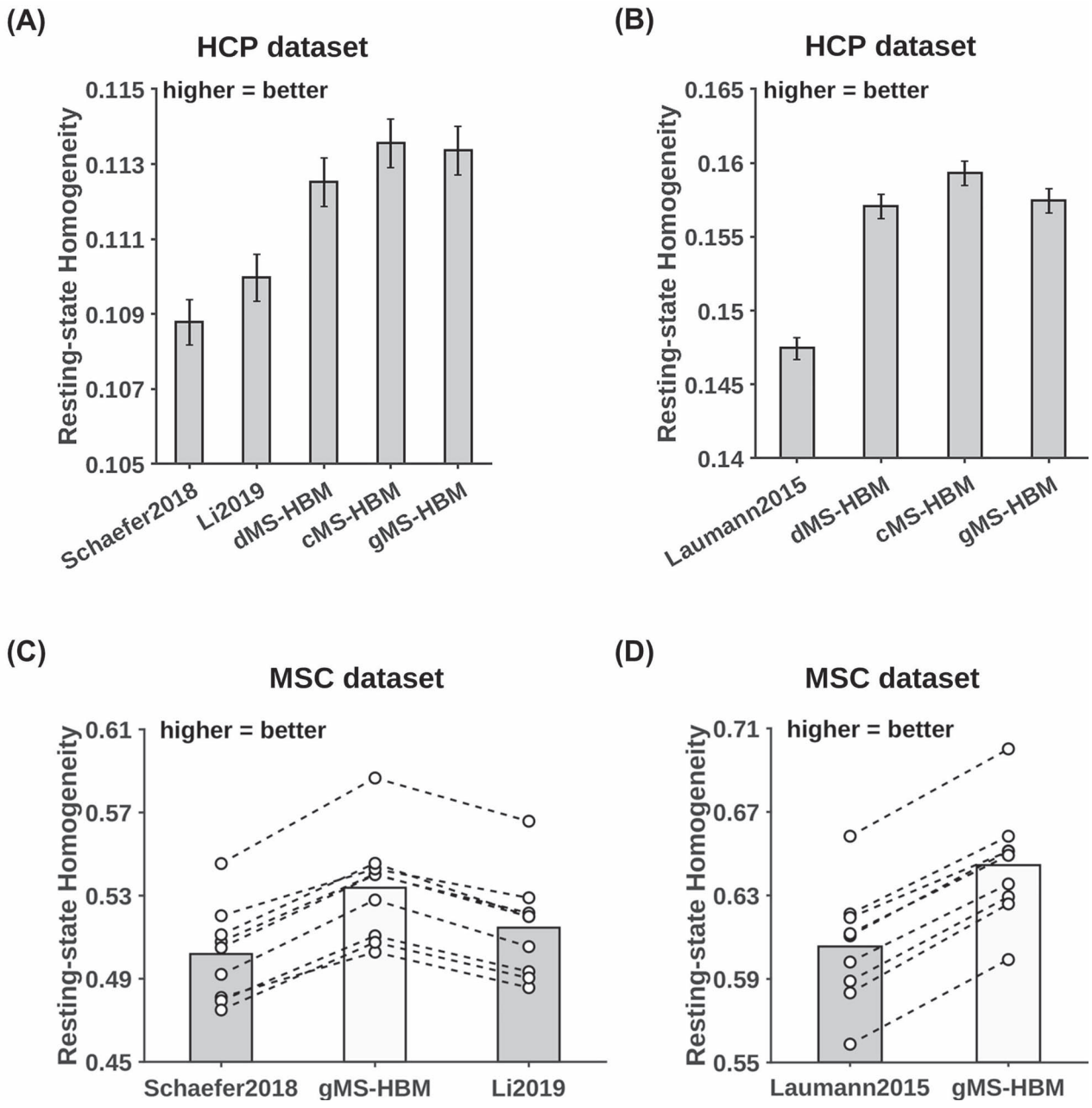

Figure 6. MS-HBM parcellations achieved better out-of-sample resting-state homogeneity than other approaches. (*A*) The 400-region individual-specific parcellations were estimated using a single rs-fMRI session and resting-state homogeneity was computed on the remaining sessions for each HCP test participant. Error bars correspond to standard errors. (*B*) Same as (*A*) except that Laumann2015 allowed different number of parcels across participants, so we matched the number of MS-HBM parcels to Laumann2015 for each participant. Therefore, the numbers for (*A*) and (*B*) were not comparable. (*C*) The 400-region individual-specific parcellations were estimated using a single rs-fMRI session and resting-state homogeneity was computed on the remaining sessions for each MSC participant. Each circle represents one MSC participant. Dash lines connect the same participants. (*D*) Same as (*C*) except that Laumann2015 allowed different number of parcels across participants, so we matched the number of MS-HBM parcels to Laumann2015 for each participant. Results for dMS-HBM and cMS-HBM in the MSC dataset are shown in Supplementary Figure S9.

Supplementary Tables S2 and S3 summarize the average prediction accuracies (Pearson's correlation) for different sets of behavioral measures, including cognitive, personality, and emotion measures. Overall, individual-specific functional connectivity strength from MS-HBM parcellations achieved better prediction performance than other approaches. In general, gMS-HBM achieved better prediction performance than dMS-HBM and cMS-HBM, but differences were not significant.

Figure 9A shows the average prediction accuracies of all 58 behaviors across different parcellation approaches. Compared with Schaefer2018 and Li2019, gMS-HBM achieved improvements of 16% ($P = 5.0e{-4}$) and 18% ($P = 5.4e{-4}$), respectively.

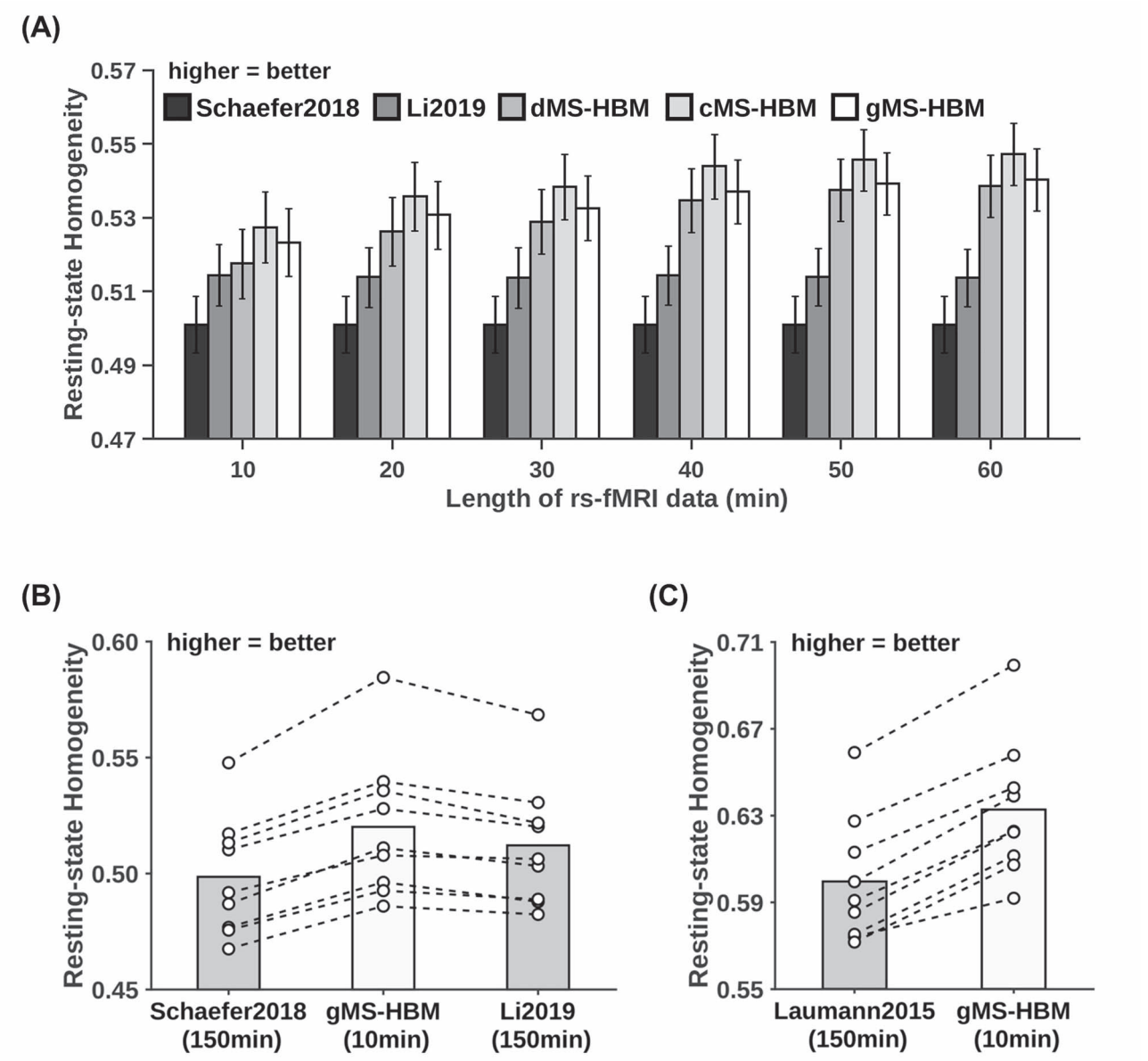

Figure 7. MS-HBM parcellations achieved better out-of-sample resting-state homogeneity with less amount of data. (*A*) The 400-region individual-specific parcellations were estimated using different lengths of rs-fMRI data for each MSC participant. Resting-state homogeneity was evaluated using leave-out sessions. Error bars correspond to standard errors. (*B*) The 400-region individual-specific parcellations were estimated for each MSC participant using 10 min of rs-fMRI data for gMS-HBM and 150 min of rs-fMRI data for Li2019. Each circle represents one MSC participant. Dash lines connect the same participants. (*C*) Same as (*B*) except that Laumann2015 yielded different number of parcels for each participant, so we matched the number of MS-HBM parcels accordingly for each participant. Results for dMS-HBM and cMS-HBM are shown in Supplementary Figure S10.

Both *P* values remained significant after correcting for multiple comparisons with FDR $q < 0.05$. Compared with cMS-HBM and dMS-HBM, gMS-HBM achieved an improvement of 5.5% and 3.4%, respectively. However, differences among MS-HBM variants were not significant.

We note that some behavioral measures were predicted poorly by all approaches. This is not unexpected because we do not expect all behavioral measures to be predictable with RSFC. Therefore, we further consider a subset of behavioral measures that could be predicted well by at least one approach. Figure 9B shows the average prediction accuracies of 36 behaviors with accuracies higher than 0.1 for at least one approach ("36 behaviors >0.1"). Compared with Schaefer2018 and Li2019, gMS-HBM achieved improvements of 13% ($P = 2.2e{-}4$) and 13% ($P = 4.5e{-}4$), respectively. All *P* values remained significant after correcting for multiple comparisons with FDR $q < 0.05$. Differences among MS-HBM variants were again not significant. Similar conclusions were obtained with COD instead of correlations (Fig. 10 and Supplementary Tables S4 and S5).

Task Performance Measures Are More Predictable Than Self-Reported Measures

To explore which behavioral measures can be consistently predicted well regardless of parcellations, we ordered the 58 behavioral measures based on averaged prediction accuracies

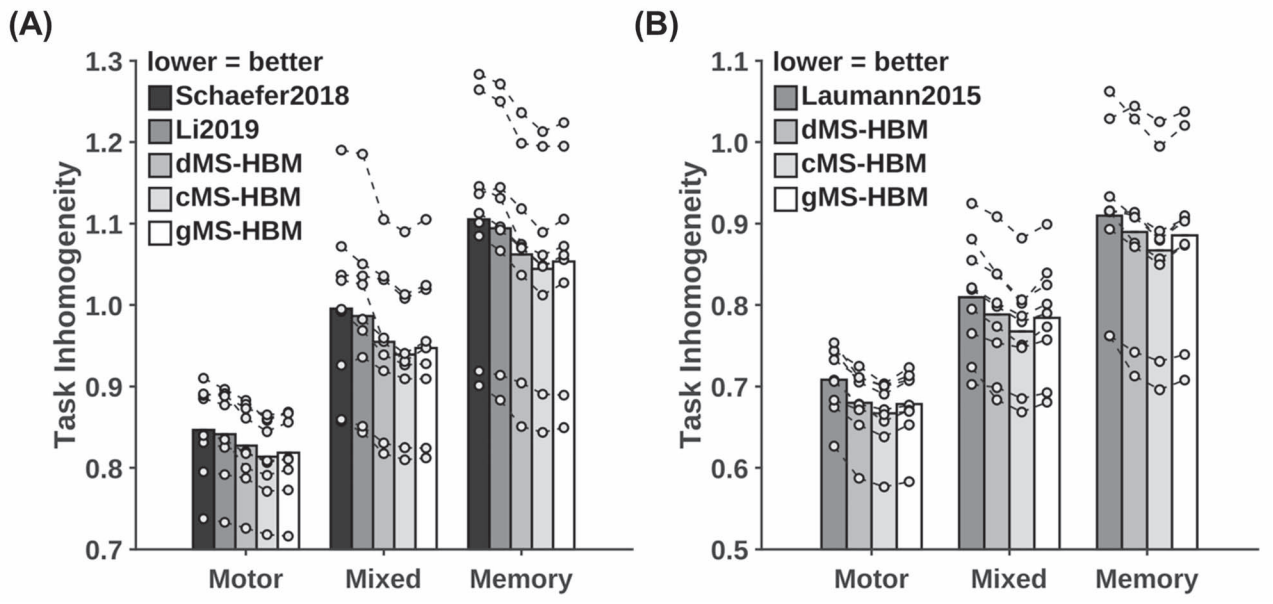

Figure 8. MS-HBM parcellations achieved better task inhomogeneity in the MSC dataset. (*A*) The 400-region individual-specific parcellations were estimated using all resting-state fMRI sessions. Task inhomogeneity was evaluated using task fMRI. Task inhomogeneity was then defined as the SD of task activation within each parcel and then averaged across all parcels and contrasts within each behavioral domain. Lower value indicates better task inhomogeneity. Each circle represents one MSC participant. Dash lines connect the same participants. (*B*) Same as (*A*) except that Laumann2015 yielded different number of parcels for each participant, so we matched the number of MS-HBM parcels accordingly for each participant. HCP results are shown in Supplementary Figure S11.

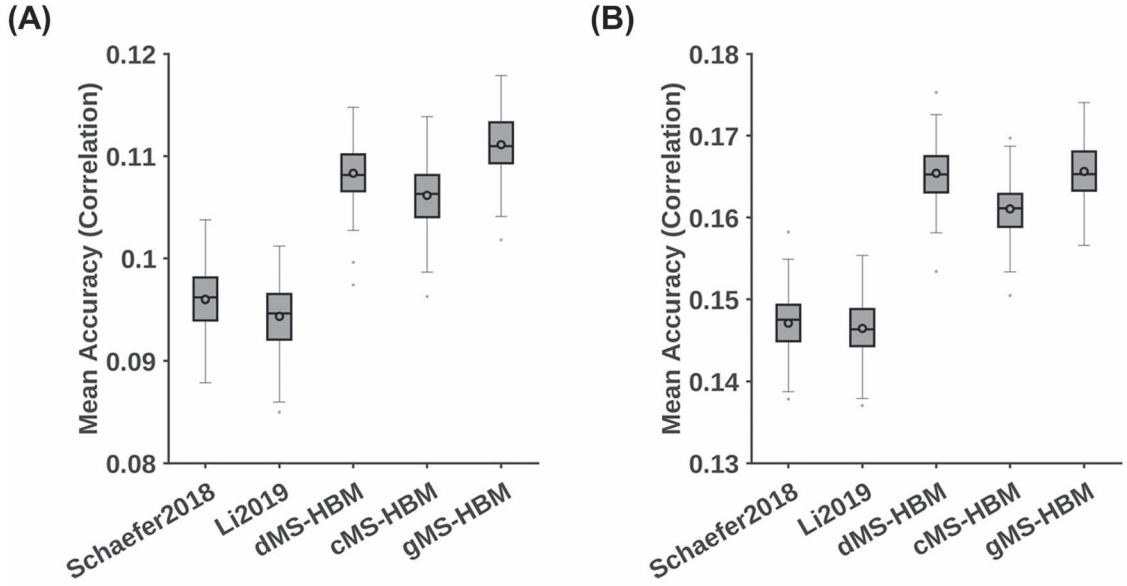

Figure 9. MS-HBM achieves the best behavioral prediction performance as measured by Pearson's correlation. (*A*) Average prediction accuracies (Pearson's correlation) of all 58 behavioral measures. Boxplots utilized default Matlab parameters, that is, box shows median and interquartile range (IQR). Whiskers indicate 1.5 IQR (not SD). Circle indicates mean. dMS-HBM, cMS-HBM, and gMS-HBM achieved average prediction accuracies of $r = 0.1083 \pm 0.0031$ (mean $\pm$ SD), 0.1062 ± 0.0031, and 0.1111 ± 0.0031, respectively. On the other hand, Schaefer2018 and Li2019 achieved average prediction accuracies of $r = 0.0960 \pm 0.0031$ and 0.0944 ± 0.0031, respectively. (*B*) Average prediction accuracies (Pearson's correlation) of 36 behavioral measures with accuracies (Pearson's correlation) higher than 0.1 for at least one approach ("36 behaviors > 0.1"). dMS-HBM, cMS-HBM, and gMS-HBM achieved average prediction accuracies of $r = 0.1630 \pm 0.0034$ (mean $\pm$ SD), 0.1590 ± 0.0035, and 0.1656 ± 0.0036, respectively. On the other hand, Schaefer2018 and Li2019 achieved average prediction accuracies of $r = 0.1442 \pm 0.0036$ and 0.1444 ± 0.0035, respectively.

(Pearson's correlation) across Schaefer2018, Li2019, and the three MS-HBM variants (Fig. 11B). Our previous studies (Liégeois et al. 2019; Li, Kong, et al. 2019a) have suggested that "self-reported" and "task performance" measures might be differentially predicted under different conditions. Using the same classification of behavioral measures (Liégeois et al. 2019; Li, Kong, et al. 2019a), we found that the average prediction accuracies of self-reported measures and task performance measures were

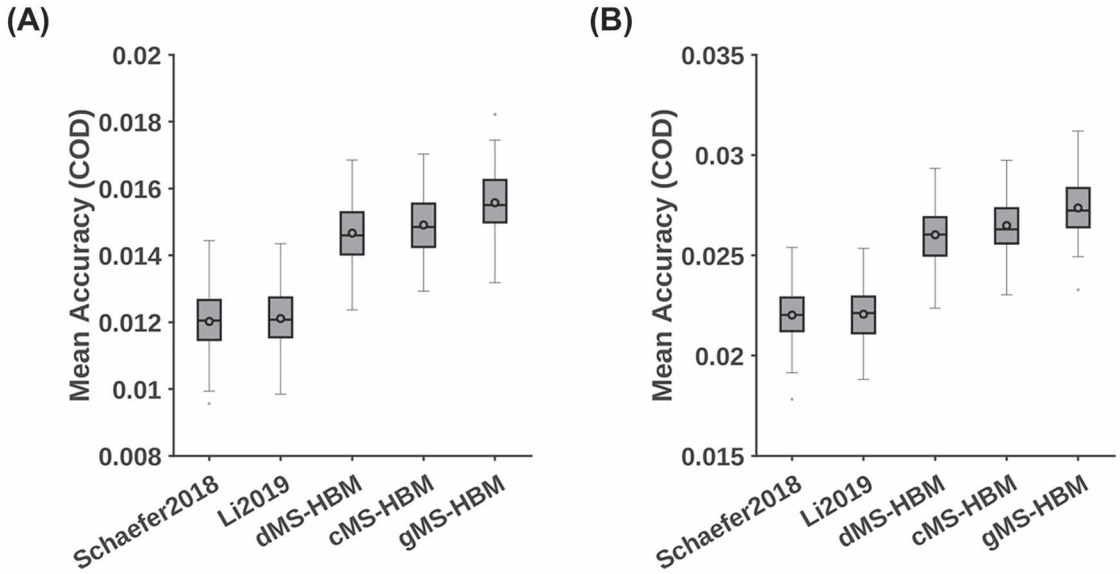

Figure 10. MS-HBM achieves the best behavioral prediction performance as measured by COD. (*A*) Average prediction accuracies (COD) of all 58 behavioral measures. Boxplots utilized default Matlab parameters, that is, box shows median and interquartile range (IQR). Whiskers indicate 1.5 IQR (not SD). Circle indicates mean. dMS-HBM, cMS-HBM, and gMS-HBM achieved average prediction accuracies (COD) = 0.0147 ± 0.0009 (mean ± SD), 0.0149 ± 0.0009, and 0.0156 ± 0.0010, respectively. On the other hand, Schaefer2018 and Li2019 achieved average prediction accuracies (COD) = 0.0120 ± 0.0009 and 0.0121 ± 0.0009, respectively. (*B*) Average prediction accuracies (COD) of 36 behavioral measures with accuracies (Pearson's correlation) greater than 0.1 for at least one approach ("36 behaviors > 0.1"). dMS-HBM, cMS-HBM, and gMS-HBM achieved average prediction accuracies (COD) = 0.0252 ± 0.0014 (mean ± SD), 0.0257 ± 0.0014, and 0.0266 ± 0.0014, respectively. On the other hand, Schaefer2018 and Li2019 achieved average prediction accuracies (COD) = 0.0212 ± 0.0014 and 0.0213 ± 0.0014, respectively.

$r = 0.0890 \pm 0.0048$ and $r = 0.1181 \pm 0.0033$, respectively (Fig. 11A), suggesting that on average, task performance measures were more predictable than self-reported measures ($P = 0.042$).

Discussion

In this manuscript, we demonstrated the robustness of the MS-HBM areal-level parcellation approach. Compared with a group-level parcellation and two state-of-the-art individual-specific areal-level parcellation approaches, we found that MS-HBM parcels were more homogeneous during resting-state while also exhibiting more uniform task activation patterns (i.e., lower task inhomogeneity). Furthermore, RSFC derived from individual-specific MS-HBM parcellations achieved better behavioral prediction performance than other approaches. Among the three MS-HBM variants, the cMS-HBM exhibited the best resting homogeneity and task inhomogeneity, while the gMS-HBM exhibited the best behavioral prediction performance.

Interpretation of the MS-HBM Areal-Level Parcellations

Previous studies have estimated around 300–400 classically defined cortical areas in the human cerebral cortex (Van Essen, Glasser, et al. 2012b). Therefore, various groups (including ours) have most frequently utilized the 400-region Schaefer group-level parcellation (Varikuti et al. 2018; Franzmeier et al. 2019; Kebets et al. 2019; Murphy et al. 2020; Orban et al. 2020). Other studies have opted to utilize different resolutions of the Schaefer group-level parcellation, for example, 100 regions (Chin Fatt et al. 2019), 200 regions (Anderson et al. 2020; Faskowitz et al. 2020), and 800 regions (Valk et al. 2020). Despite our focus on the 400-region areal-level parcellations in the current study, we do not believe that there is an optimal number of cortical parcels because of the multi-resolution organization of the cerebral cortex (Churchland and Sejnowski 1988; van den Heuvel and Yeo 2017). Indeed, given the heterogeneity of cortical areas (Kaas 1987; Amunts and Zilles 2015), cortical areas might be further subdivided into meaningful computational subunits.

More specifically and consistent with other studies, our areal-level parcels likely captured subareal features such as somatotopy and visual eccentricity (Gordon et al. 2016; Schaefer et al. 2018). Ultimately, the choice of parcellation resolution might depend on the specific application. For example, a recent study suggested that brain–behavior relationships are scale-dependent (Betzel et al. 2019). Furthermore, a higher resolution parcellation might be computationally infeasible for certain analysis, such as edge-centric network analysis (Faskowitz et al. 2020). Therefore, we have provided trained MS-HBM at different spatial resolutions, ranging from 100 to 1000 parcels. It is worth noting that because our parcels do not correspond to traditional cortical areas (Kaas 1987; Amunts and Zilles 2015), we have been careful to avoid the term "areas." Instead we use the term "areal-level parcellation" when referring to the entire parcellation and "parcels" when referring to individual regions throughout the manuscript.

Several studies have shown that brain networks reconfigure during tasks (Cole et al. 2014; Krienen et al. 2014; Salehi et al. 2019). Consequently, some have questioned the existence of a single individual-specific areal-level parcellation that generalizes across resting and task states (Salehi et al. 2019). While we do not contest the results of Salehi and colleagues, we have a very different interpretation. Cortical areas (e.g., V1) are conceptualized as representing stable computational units (Felleman and Van Essen 1991). Consequently, their boundaries should remain the same regardless of transient task states across the span of a few days, even if long-term experiences can potentially

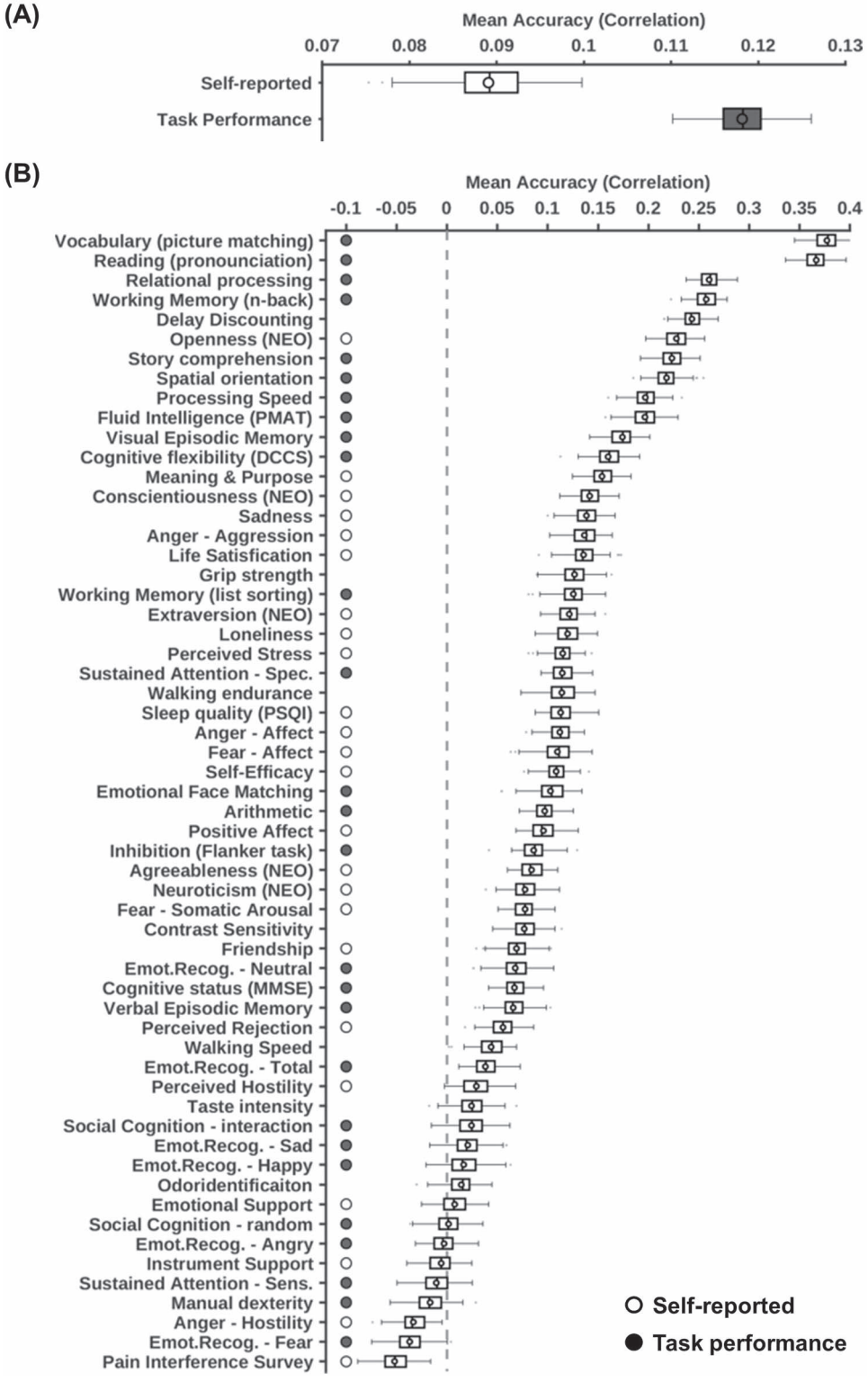

Figure 11. Task performance measures were predicted better than self-reported measures across different parcellation approaches. Prediction accuracies were averaged across all parcellation approaches (three MS-HBM variants, Schaefer2018, and Li2019). (A) Prediction accuracies averaged across HCP task-performance measures (gray) and HCP self-reported measures (white). (B) Behavioral measures were ordered based on average prediction accuracies. Gray color indicates task performance measures. White color indicates self-reported measures. Boxplots utilized default Matlab parameters, that is, box shows median and interquartile range (IQR). Whiskers indicate 1.5 IQR (not SD). Circle indicates mean. Designation of behavioral measures into "self-reported" and "task-performance" measures followed previous studies (Liégeois et al. 2019; Li et al. 2019a).

shape the development and formation of cortical areas (Arcaro et al. 2017; Gomez et al. 2019). Thus, the results of Salehi and colleagues do not rule out the plausibility of estimating a stable individual-level areal-level parcellation with rs-fMRI data alone. Rather, Salehi and colleagues motivate the need to estimate areal-level parcellations jointly from rs-fMRI, task-fMRI, and other modalities (Glasser et al. 2016; Eickhoff, Constable, et al. 2018a) in order to achieve invariance across brain states. We leave this for future work.

MS-HBM Areal-Level Parcellations Are More Homogeneous Than Other Approaches in Out-of-Sample Resting- and Task-fMRI

Dealing with RSFC matrices at the original voxel or vertex resolution is difficult because of the high-dimensionality. Thus, areal-level brain parcellations have been widely utilized as a dimensionality reduction tool (Eickhoff, Constable, et al. 2018a), for example, averaged time course of a parcel is used to represent the entire parcel (Varoquaux and Craddock 2013; Finn et al. 2015; Rosenberg et al. 2016). For the dimensionality reduction to be valid, vertices within each areal-level parcel should have similar time courses, that is, high resting-state homogeneity (Gordon et al. 2016; Schaefer et al. 2018). Across two datasets (HCP and MSC), we found that MS-HBM areal-level parcellations exhibited higher resting-state homogeneity than three other approaches, suggesting that rs-fMRI time courses are more similar within MS-HBM parcels (Figs 6 and 7).

Furthermore, if an individual-specific areal-level parcellation accurately captures the functional brain organization of a participant, one might expect task activation to be uniform within parcels, that is, low task inhomogeneity (Gordon, Laumann, Gilmore, et al. 2017b; Schaefer et al. 2018). We found that MS-HBM parcellations achieved better task inhomogeneity than other approaches in both HCP and MSC datasets (Fig. 8 and Supplementary Fig. S11). Given the strong link between task-fMRI and rs-fMRI (Smith et al. 2009; Mennes et al. 2010; Cole et al. 2014; Krienen et al. 2014; Bertolero et al. 2015; Yeo et al. 2015; Tavor et al. 2016), this is perhaps not surprising.

Intriguingly, the improvement in task inhomogeneity varied significantly across task domains (Fig. 8 and Supplementary Fig. S11) with the motor task exhibiting the least task inhomogeneity improvement for both HCP and MSC datasets. Given that the motor domain exhibited one of the lowest task inhomogeneity across behavioral domains, there might not be much room for improvement. Furthermore, sensory-motor parcels exhibited low inter-subject variation in terms of location and spatial topography (Fig. 4B), so different approaches might perform similarly well. Other possible reasons might include variation in task design and duration.

It is worth pointing out that even though MSC dataset only contained nine participants, MS-HBM parcellations exhibited better resting homogeneity and task inhomogeneity in every single participant (Figs 6–8 and Supplementary Figures S9–S11). This suggests that MS-HBM parameters estimated from HCP were effective in MSC despite significant acquisition and preprocessing differences.

MS-HBM Works Well Even with Only 10 min of rs-fMRI Data

It is well known that longer scan durations can improve the reliability of RSFC measures (Van Dijk et al. 2010; Xu et al. 2016; Kong et al. 2019). Recent studies have suggested that at least 20–30 min of data is needed to obtain reliable measurements (Laumann et al. 2015; O'Connor et al. 2017; Gordon, Laumann, Gilmore, et al. 2017b). Consistent with previous work, we found that resting-state homogeneity of individual-specific areal-level parcellations continued to improve with more data (Fig. 7 and Supplementary Fig. S10). The improvements started to plateau around 40–50 min of data.

Although MS-HBM required multi-session rs-fMRI data for training, the models could be applied to a single rs-fMRI session from a new dataset. More specifically, in the MSC dataset, we showed that MS-HBM areal-level parcellations estimated with only 10 min of rs-fMRI data exhibited better resting-state homogeneity than two other approaches using 150 min of data (Gordon, Laumann, Gilmore, et al. 2017b; Li, Wang, et al. 2019b).

RSFC of Individual-Specific MS-HBM Parcellations Improves Behavioral Prediction

A vast body of literature has shown that functional connectivity derived from group-level parcellations can be utilized for behavioral prediction (Hampson et al. 2006; Finn et al. 2015; Smith et al. 2015; Yeo et al. 2015; Rosenberg et al. 2016; He et al. 2020). However, there is a preponderance of evidence that group-level parcellations obscure individual-specific topographic features (Harrison et al. 2015; Laumann et al. 2015; Langs et al. 2016; Braga and Buckner 2017; Chong et al. 2017; Gordon, Laumann, Adeyemo, et al. 2017a, Gordon, Laumann, Gilmore, 2017b), which are behaviorally meaningful (Bijsterbosch et al. 2018, 2019; Kong et al. 2019; Seitzman et al. 2019). Recent studies have also suggested that functional connectivity strength derived from individual-specific parcellations might also improve behavioral prediction (Pervaiz et al. 2019; Li, Wang, et al. 2019b).

We found that MS-HBM parcellations captured individual-specific features that were replicable across sessions (Fig. 5 and Supplementary Figs S2–S7). Furthermore, RSFC derived from individual-specific MS-HBM areal-level parcellations achieved better behavioral prediction performance compared with a group-level parcellation (Schaefer et al. 2018) and a recently published individual-specific parcellation approach (Li, Wang, et al. 2019b). Overall, our results suggest that individual differences in functional connectivity strength of MS-HBM parcels were more behaviorally meaningful than of other parcellation approaches.

It is worth noting that the absolute improvement in prediction performance was modest on average, although some behavioral measures appeared to benefit more than others. For example, when comparing Li2019 and gMS-HBM for behavioral prediction, the prediction accuracy (Pearson's correlation) of "openness (NEO)" improved from 0.19 to 0.26, while the accuracy (Pearson's correlation) of "vocabulary (picture matching)" improved from 0.36 to 0.39. Thus, gMS-HBM might be more helpful for predicting certain behavioral measures than others.

Further analysis suggested that task performance measures were on average predicted with higher accuracy than self-reported measures (Fig. 11). This differentiation between task performance and self-reported measures was consistent with previous investigations of RSFC–behavior relationships. For example, RSFC has been shown to predict cognition better than personality and mental health (Dubois et al. 2018; Chen et al. 2020). Dynamic functional connectivity is also more strongly associated with cognition and task performance than self-reported measures (Vidaurre et al. 2017; Liégeois et al.

2019). Finally, regressing the global signal has been shown to improve the prediction of task performance measures more than self-reported measures (Li, Kong, et al. 2019a).

Spatially Localized Individual-Specific Areal-Level Parcels

Postmortem studies have generally identified cortical areas that are spatially contiguous (Kaas 1987; Felleman and Van Essen 1991; Amunts and Zilles 2015). This has motivated most resting-state areal-level parcellations to estimate spatially contiguous parcels (Shen et al. 2013; Honnorat et al. 2015; Gordon et al. 2016; Chong et al. 2017). One approach to achieve spatially contiguous parcels is to introduce a spatial connectedness term into the optimization objective so that distributed parcels would have large penalty (Honnorat et al. 2015; Schaefer et al. 2018). Another approach is to start with initial spatially contiguous parcels and to iteratively adjust the boundaries to maintain spatial contiguity (Blumensath et al. 2012; Chong et al. 2017; Salehi et al. 2019). Yet another method is to utilize the local-gradient approach, which detects sharp transitions in RSFC profiles, followed by a postprocessing procedure (Cohen et al. 2008; Gordon et al. 2016). However, work from Glasser et al. (2016) suggested that some individual-specific areal-level parcels might comprise multiple spatially close components in some individuals.

Given the lack of consensus, we explored three MS-HBM variants in this study. We found that strictly contiguous cMS-HBM parcels achieved the best out-of-sample resting-state homogeneity and task inhomogeneity (Figs 6–8 and Supplementary Figs S8–S10). One possible reason is that cMS-HBM parcellation boundaries were smoother than dMS-HBM and gMS-HBM parcellations. Since fMRI data are spatially smooth, parcellations with smoother boundaries might have an inherent homogeneity advantage, without necessarily being better at capturing true areal boundaries. Another potential artifact of smooth data is the appearance of excessively round parcels that are at odds with histological studies, which show that cortical areas express diverse spatial configurations.

Based on our geometric analyses, we found gMS-HBM to be most anatomically plausible among the three parcellations, having both fewer spatially disconnected components than dMS-HBM, and intermediate levels of roundness between dMS-HBM and cMS-HBM. Furthermore, RSFC derived from gMS-HBM parcels achieved the best behavioral prediction performance, albeit not reaching statistical significance (Fig. 9 and Supplementary Figure S11; Supplementary Tables S2–S5). As elaborated in previous studies (Gordon et al. 2016; Schaefer et al. 2018; Kong et al. 2019), assessment of parcellations should integrate and weigh performance across multiple metrics. For the reasons outlined above, we prefer individual-specific gMS-HBM areal-level parcellations among the three MS-HBM variants.

Overall, our findings suggest that the brain's large-scale organization might potentially comprise certain functional regions that are spatially disconnected. Neuronal migration, guided by cell-to-cell interactions and gradients of diffusible cues, plays an important role in establishing the brain's complex cytoarchitectonic organization during embryogenesis (Silva et al. 2019). Spatially disconnected parcels might reflect functionally analogous neuronal populations from the same cellular lineage that separate due to natural variation in migration patterns in early development.

That said, we are aware that one cannot establish with certainty the existence of spatially disconnected cortical areas based on resting-fMRI data alone. It is possible that disconnected components of a noncontiguous parcel are inseparable by resting-fMRI measurements but are separable by other neural properties, such as microstructure or task activations. Given that fMRI is an indirect measurement of neuronal signals, the functional coupling among disconnected components could also be driven by non-neural mechanisms (e.g., vasculature).

Nevertheless, our individual-level areal parcellation provides an explicit model that can be further validated using prospectively acquired rs-fMRI paired with other approaches, for example, post-mortem histological analyses (Xu et al. 2018; Hayashi et al. 2020) or with spatially targeted intracranial recording (Wang et al. 2015; Fox et al. 2018).

Conclusions

We proposed a MS-HBM that accounted for both inter-subject and intra-subject functional connectivity variability when estimating individual-specific areal-level parcellations. Three MS-HBM variants with different spatial localization priors were explored. Using 10 min of rs-fMRI data, individual-specific MS-HBM areal-level parcellations generalized better to out-of-sample rs-fMRI data from the same participants than a group-level parcellation approach and two prominent individual-specific areal-level parcellation approaches using 150 min of rs-fMRI data. Furthermore, RSFC derived from MS-HBM parcellations exhibited better behavioral prediction performance than alternative parcellation approaches.

Supplementary Material

Supplementary material is available at *Cerebral Cortex* online.

Notes

We like to thank Danilo Bzdok, John Murray, Michael Breakspear, Peter Fox, and Jack Lancaster for insightful comments and feedback on this work. *Conflict of Interest*: None declared.

Funding

This work was supported by the Singapore National Research Foundation (NRF) Fellowship (Class of 2017), the Singapore Ministry of Defense (Project CURATE), and the National University of Singapore Yong Loo Lin School of Medicine (NUHSRO/2020/124/TMR/LOA). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not reflect the views of the Singapore Ministry of Defense and National Research Foundation, Singapore. Computational work for this article was partially performed on resources of the National Supercomputing Centre, Singapore (https://www.nscc.sg). 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 in part 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, Zilles K. 2015. Architectonic mapping of the human brain beyond Brodmann. *Neuron*. 88:1086–1107.

Anderson KM, Collins MA, Kong R, Fang K, Li J, He T, Chekroud AM, Yeo BTT, Holmes AJ. 2020. Convergent molecular, cellular, and cortical neuroimaging signatures of major depressive disorder. *PNAS*. 117:25138–25149.

Arcaro MJ, Schade PF, Vincent JL, Ponce CR, Livingstone MS. 2017. Seeing faces is necessary for face-domain formation. *Nat Neurosci*. 20:1404–1412.

Arslan S, Parisot S, Rueckert D. 2015. Joint spectral decomposition for the parcellation of the human cerebral cortex using resting-state fMRI. In: Ourselin S, Alexander DC, Westin C-F, Cardoso MJ, editors. *Information processing in medical imaging*. Springer International Publishing: Cham, pp. 85–97.

Barch DM, Burgess GC, Harms MP, Petersen SE, Schlaggar BL, Corbetta M, Glasser MF, Curtiss S, Dixit S, Feldt C, et al. 2013. Function in the human connectome: task-fMRI and individual differences in behavior. *Neuroimage*. 80:169–189.

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.

Betzel RF, Bertolero MA, Gordon EM, Gratton C, Dosenbach NUF, Bassett DS. 2019. The community structure of functional brain networks exhibits scale-specific patterns of inter- and intra-subject variability. *Neuro Image*. 202:115990.

Bijsterbosch JD, Beckmann CF, Woolrich MW, Smith SM, Harrison SJ. 2019. The relationship between spatial configuration and functional connectivity of brain regions revisited. *Elife*. 8:e44890.

Bijsterbosch JD, Woolrich MW, Glasser MF, Robinson EC, Beckmann CF, Van Essen DC, Harrison SJ, Smith SM. 2018. The relationship between spatial configuration and functional connectivity of brain regions. *Elife*. 7:e32992.

Birn RM, Molloy EK, Patriat R, Parker T, Meier TB, Kirk GR, Nair VA, Meyerand ME, Prabhakaran V. 2013. The effect of scan length on the reliability of resting-state fMRI connectivity estimates. *Neuro Image*. 83:550–558.

Biswal B, Yetkin FZ, 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.

Blumensath T, Jbabdi S, Glasser MF, Van Essen DC, Ugurbil K, Behrens TEJ, Smith SM. 2013. Spatially constrained hierarchical parcellation of the brain with resting-state fMRI. *Neuroimage*. 76:313–324.

Blumensath T, Behrens TEJ, Smith SM. 2012. Resting-state FMRI single subject cortical parcellation based on region growing. *Med Image Comput Comput Assist Interv*. 15:188–195.

Bouckaert RR, Frank E. 2004. Evaluating the replicability of significance tests for comparing learning algorithms. In: Dai H, Srikant R, Zhang C, editors. *Advances in knowledge discovery and data mining*. Springer: Berlin, Heidelberg, pp. 3–12.

Braga RM, Buckner RL. 2017. Parallel interdigitated distributed networks within the individual estimated by intrinsic functional connectivity. *Neuron*. 95, e5:457–471.

Bressler SL, Menon V. 2010. Large-scale brain networks in cognition: emerging methods and principles. *Trends Cogn Sci*. 14:277–290.

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

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 Connect*. 6:669–680.

Chen B, Xu T, Zhou C, Wang L, Yang N, Wang Z, Dong H-M, Yang Z, Zang Y-F, Zuo X-N, et al. 2015. Individual variability and test-retest reliability revealed by ten repeated resting-state brain scans over one month. *PLoS One*. 10:e0144963.

Chen J, Tam A, Kebets V, Orban C, Ooi LQR, Marek S, Dosenbach N, Eickhoff S, Bzdok D, Holmes AJ, et al. 2020. Shared and unique brain network features predict cognition, personality and mental health in childhood. *bioRxiv*. 2020.06.24.168724.

Chin Fatt CR, Jha MK, Cooper CM, Fonzo G, South C, Grannemann B, Carmody T, Greer TL, Kurian B, Fava M, et al. 2019. Effect of intrinsic patterns of functional brain connectivity in moderating antidepressant treatment response in major depression. *AJP*. 177:143–154.

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. *Neuro Image*. 156:87–100.

Churchland PS, Sejnowski TJ. 1988. Perspectives on cognitive neuroscience. *Science*. 242:741–745.

Cohen AL, Fair DA, Dosenbach NUF, Miezin FM, Dierker D, Van Essen DC, Schlaggar BL, Petersen SE. 2008. Defining functional areas in individual human brains using resting functional connectivity MRI. *Neuroimage*. 41:45–57.

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.

Craddock RC, James GA, Holtzheimer PE, Hu XP, Mayberg HS. 2012. A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Hum Brain Mapp*. 33: 1914–1928.

Cui Z, Li H, Xia CH, Larsen B, Adebimpe A, Baum GL, Cieslak M, Gur RE, Gur RC, Moore TM, et al. 2020. Individual variation in control network topography supports executive function in youth. *Neuron*. 106:340–353.e8.

Destrieux C, Fischl B, Dale A, Halgren E. 2010. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuro Image*. 53:1–15.

Dosenbach NUF, Nardos B, Cohen AL, Fair DA, Power JD, Church JA, Nelson SM, Wig GS, Vogel AC, Lessov-Schlaggar CN, et al. 2010. Prediction of individual brain maturity using fMRI. *Science*. 329:1358–1361.

Dubois J, Galdi P, Paul LK, Adolphs R. 2018. A distributed brain network predicts general intelligence from resting-state human neuroimaging data. *Philos Trans R Soc Lond B Biol Sci*. 373.

Eickhoff SB, Constable RT, Yeo BTT. 2018a. Topographic organization of the cerebral cortex and brain cartography. *Neuro Image*. 170:332–347.

Eickhoff SB, Yeo BTT, Genon S. 2018b. Imaging-based parcellations of the human brain. *Nat Rev Neurosci*. 19:672–686.

Faskowitz J, Esfahlani FZ, Jo Y, Sporns O, Betzel RF. 2020. Edge-centric functional network representations of human cerebral cortex reveal overlapping system-level architecture. *Nat Neurosci*. 23:1644–1654.

Felleman DJ, Van Essen DC. 1991. Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex*. 1:1–47.

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, 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.

Fox MD, Raichle ME. 2007. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci*. 8:700–711.

Fox KCR, Foster BL, Kucyi A, Daitch AL, Parvizi J. 2018. Intracranial electrophysiology of the human default network. *Trends Cogn Sci*. 22:307–324.

Franzmeier N, Rubinski A, Neitzel J, Kim Y, Damm A, Na DL, Kim HJ, Lyoo CH, Cho H, Finsterwalder S, et al. 2019. Functional connectivity associated with tau levels in ageing, Alzheimer's, and small vessel disease. *Brain*. 142:1093–1107.

Friston KJ, Williams S, Howard R, Frackowiak RSJ, Turner R. 1996. Movement-related effects in fMRI time-series. *Magn Reson Med*. 35:346–355.

Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, Ugurbil K, Andersson J, Beckmann CF, Jenkinson M, et al. 2016. A multi-modal parcellation of human cerebral cortex. *Nature*. 536:171–178.

Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, Xu J, Jbabdi S, Webster M, Polimeni JR, et al. 2013. The minimal preprocessing pipelines for the Human Connectome Project. *Neuro Image*. 80:105–124.

Goldman-Rakic PS. 1988. Topography of cognition: parallel distributed networks in primate association cortex. *Annu Rev Neurosci*. 11:137–156.

Gomez J, Barnett M, Grill-Spector K. 2019. Extensive childhood experience with Pokémon suggests eccentricity drives organization of visual cortex. *Nat Hum Behav*. 3:611–624.

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, Gilmore AW, Nelson SM, Dosenbach NUF, Petersen SE. 2017a. Individual-specific features of brain systems identified with resting state functional correlations. *Neuro Image*. 146:918–939.

Gordon EM, Laumann TO, Gilmore AW, Newbold DJ, Greene DJ, Berg JJ, Ortega M, Hoyt-Drazen C, Gratton C, Sun H, et al. 2017b. Precision functional mapping of individual human brains. *Neuron*. 95, e7:791–807.

Gratton C, Laumann TO, Nielsen AN, Greene DJ, Gordon EM, Gilmore AW, Nelson SM, Coalson RS, Snyder AZ, Schlaggar BL, et al. 2018. Functional brain networks are dominated by stable group and individual factors, not cognitive or daily variation. *Neuron*. 98, e5:439–452.

Griffanti L, Salimi-Khorshidi G, Beckmann CF, Auerbach EJ, Douaud G, Sexton CE, Zsoldos E, Ebmeier KP, Filippini N, Mackay CE, et al. 2014. ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *Neuroimage*. 95:232–247.

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.

Hayashi T, Hou Y, Glasser MF, Autio JA, Knoblauch K, Inoue-Murayama M, Coalson T, Yacoub E, Smith S, Kennedy H, et al. 2020. The nonhuman primate neuroimaging & neuroanatomy project. *arXiv*. 2010.00308 [q-Bio].

He T, Kong R, Holmes AJ, Nguyen M, Sabuncu MR, Eickhoff SB, Bzdok D, Feng J, Yeo BTT. 2020. Deep neural networks and kernel regression achieve comparable accuracies for functional connectivity prediction of behavior and demographics. *Neuro Image*. 206:116276.

van den Heuvel MP, Yeo BTT. 2017. A spotlight on bridging microscale and macroscale human brain architecture. *Neuron*. 93:1248–1251.

Honnorat N, Eavani H, Satterthwaite TD, Gur RE, Gur RC, Davatzikos C. 2015. GraSP: geodesic graph-based segmentation with shape priors for the functional parcellation of the cortex. *Neuro Image*. 106:207–221.

Kaas JH. 1987. The organization of neocortex in mammals: implications for theories of brain function. *Annu Rev Psychol*. 38:129–151.

Kebets V, Holmes AJ, Orban C, Tang S, Li J, Sun N, Kong R, Poldrack RA, Yeo BTT. 2019. Somatosensory-motor dysconnectivity spans multiple transdiagnostic dimensions of psychopathology. *Biol Psychiatry*. 86:779–791.

Kong R, Li J, Orban C, Sabuncu MR, Liu H, Schaefer A, Sun N, Zuo X-N, Holmes AJ, Eickhoff SB, et al. 2019. Spatial topography of individual-specific cortical networks predicts human cognition, personality, and emotion. *Cereb Cortex*. 29:2533–2551.

Krienen FM, Yeo BTT, Buckner RL. 2014. Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture. *Philos Trans R Soc Lond B Biol Sci*. 369.

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.

Laumann TO, Gordon EM, Adeyemo B, Snyder AZ, Joo SJ, Chen M-Y, Gilmore AW, McDermott KB, Nelson SM, Dosenbach NUF, et al. 2015. Functional system and areal organization of a highly sampled individual human brain. *Neuron*. 87:657–670.

Li J, Kong R, Liégeois R, Orban C, Tan Y, Sun N, Holmes AJ, Sabuncu MR, Ge T, Yeo BTT. 2019a. Global signal regression strengthens association between resting-state functional connectivity and behavior. *Neuro Image*. 196:126–141.

Li M, Wang D, Ren J, Langs G, Stoecklein S, Brennan BP, Lu J, Chen H, Liu H. 2019b. Performing group-level functional image analyses based on homologous functional regions mapped in individuals. *PLoS Biol*. 17:e2007032.

Liégeois R, Li J, Kong R, Orban C, Van De Ville D, Ge T, Sabuncu MR, Yeo BTT. 2019. Resting brain dynamics at different timescales capture distinct aspects of human behavior. *Nat Commun*. 10:2317.

Mejia AF, Nebel MB, Shou H, Crainiceanu CM, Pekar JJ, Mostofsky S, Caffo B, Lindquist MA. 2015. Improving reliability of subject-level resting-state fMRI parcellation with shrinkage estimators. *Neuro Image*. 112:14–29.

Mejia AF, Nebel MB, Barber AD, Choe AS, Pekar JJ, Caffo BS, Lindquist MA. 2018. Improved estimation of subject-level functional connectivity using full and partial correlation with empirical Bayes shrinkage. *Neuro Image*. 172:478–491.

Mennes M, Kelly C, Zuo X-N, Di Martino A, Biswal BB, Castellanos FX, Milham MP. 2010. Inter-individual differences in resting-state functional connectivity predict task-induced BOLD activity. *Neuroimage*. 50:1690–1701.

Mesulam M-M. 1990. Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Ann Neurol*. 28:597–613.

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: a probabilistic perspective*. Cambridge (MA): MIT Press.

Murphy AC, Bertolero MA, Papadopoulos L, Lydon-Staley DM, Bassett DS. 2020. Multimodal network dynamics underpinning working memory. *Nat Commun*. 11:3035.

Mwilambwe-Tshilobo L, Ge T, Chong M, Ferguson MA, Misic B, Burrow AL, Leahy RM, Spreng RN. 2019. Loneliness and meaning in life are reflected in the intrinsic network architecture of the brain. *Soc Cogn Affect Neurosci*. 14:423–433.

Nowozin S, Lampert CH. 2010. Global interactions in random field models: a potential function ensuring connectedness. *SIAM J Imaging Sci*. 3:1048–1074.

O'Connor D, Potler NV, Kovacs M, Xu T, Ai L, Pellman J, Vanderwal T, Parra LC, Cohen S, Ghosh S, et al. 2017. The healthy brain network serial scanning initiative: a resource for evaluating inter-individual differences and their reliabilities across scan conditions and sessions. *Gigascience*. 6:giw011.

Orban C, Kong R, Li J, Chee MWL, Yeo BTT. 2020. Time of day is associated with paradoxical reductions in global signal fluctuation and functional connectivity. *PLoS Biol*. 18:e3000602.

Pervaiz U, Vidaurre D, Woolrich MW, Smith SM. 2019. Optimising network modelling methods for fMRI. *bioRxiv*. 741595.

Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, et al. 2011. Functional network organization of the human brain. *Neuron*. 72:665–678.

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.

Sabuncu MR, Thomas Yeo BT, Van Leemput K, Fischl B, Golland P. 2010. A generative model for image segmentation based on label fusion. *IEEE Trans Med Imaging*. 29:1714–1729.

Salehi M, Karbasi A, Shen X, Scheinost D, Constable RT. 2018. An exemplar-based approach to individualized parcellation reveals the need for sex specific functional networks. *Neuro Image*. 170:54–67.

Salehi M, Greene AS, Karbasi A, Shen X, Scheinost D, Constable RT. 2019. There is no single functional atlas even for a single individual: functional parcel definitions change with task. *Neuro Image*. 116366.

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. 2018. Local-global Parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb Cortex*. 28:3095–3114.

Seitzman BA, Gratton C, Laumann TO, Gordon EM, Adeyemo B, Dworetsky A, Kraus BT, Gilmore AW, Berg JJ, Ortega M, et al. 2019. Trait-like variants in human functional brain networks. *PNAS*. 116:22851–22861.

Sereno M, Dale A, Reppas J, Kwong K, Belliveau J, Brady T, Rosen B, Tootell R. 1995. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*. 268:889–893.

Shehzad Z, Kelly AMC, Reiss PT, Gee DG, Gotimer K, Uddin LQ, Lee SH, Margulies DS, Roy AK, Biswal BB, et al. 2009. The resting brain: unconstrained yet reliable. *Cereb Cortex*. 19:2209–2229.

Shen X, Tokoglu F, Papademetris X, Constable RT. 2013. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *Neuro Image*. 82:403–415.

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.

Silva CG, Peyre E, Nguyen L. 2019. Cell migration promotes dynamic cellular interactions to control cerebral cortex morphogenesis. *Nat Rev Neurosci*. 20:318–329.

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

Smith SM, Beckmann CF, Andersson J, Auerbach EJ, Bijsterbosch J, Douaud G, Duff E, Feinberg DA, Griffanti L, Harms MP, et al. 2013. Resting-state fMRI in the Human Connectome Project. *Neuro Image*. 80:144–168.

Smith SM, Nichols TE, Vidaurre D, Winkler AM, Behrens TEJ, 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.

Tavor I, Parker Jones O, Mars RB, Smith SM, Behrens TE, Jbabdi S. 2016. Task-free MRI predicts individual differences in brain activity during task performance. *Science*. 352:216–220.

Uddin LQ, Yeo BTT, Spreng RN. 2019. Towards a universal taxonomy of macro-scale functional human brain networks. *Brain Topogr*. 32:926–942.

Valk SL, Xu T, Margulies DS, Masouleh SK, Paquola C, Goulas A, Kochunov P, Smallwood J, Yeo BTT, Bernhardt BC, et al. 2020. Shaping brain structure: genetic and phylogenetic axes of macroscale organization of cortical thickness. *Sci Adv*. 6:eabb3417.

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. *J Neurophysiol*. 103:297–321.

Van Essen DC, Glasser MF. 2014. In vivo architectonics: a corticocentric perspective. *Neuroimage*. 93(Pt 2):157–164.

Van Essen DC, Ugurbil K, Auerbach E, Barch D, Behrens TEJ, Bucholz R, Chang A, Chen L, Corbetta M, Curtiss SW, et al. 2012a. The Human Connectome Project: a data acquisition perspective. *Neuroimage*. 62:2222–2231.

Van Essen DC, Glasser MF, Dierker DL, Harwell J, Coalson T. 2012b. Parcellations and hemispheric asymmetries of human cerebral cortex analyzed on surface-based atlases. *Cereb Cortex*. 22:2241–2262.

Varikuti DP, Genon S, Sotiras A, Schwender H, Hoffstaedter F, Patil KR, Jockwitz C, Caspers S, Moebus S, Amunts K, et al. 2018. Evaluation of non-negative matrix factorization of grey matter in age prediction. *Neuroimage*. 173:394–410.

Varoquaux G, Craddock RC. 2013. Learning and comparing functional connectomes across subjects. *Neuro Image*. 80:405–415.

Varoquaux G, Raamana PR, Engemann DA, Hoyos-Idrobo A, Schwartz Y, Thirion B. 2017. Assessing and tuning brain decoders: cross-validation, caveats, and guidelines. *Neuro Image*. 145:166–179.

Vidaurre D, Smith SM, Woolrich MW. 2017. Brain network dynamics are hierarchically organized in time. *Proc Natl Acad Sci*. 114:12827–12832.

Vogt B. 2009. *Cingulate Neurobiology and Disease*. New York: Oxford University Press.

Wang D, Buckner RL, Fox MD, Holt DJ, Holmes AJ, Stoecklein S, Langs G, Pan R, Qian T, Li K, et al. 2015. Parcellating cortical functional networks in individuals. *Nat Neurosci*. 18:1853–1860.

Weis S, Patil KR, Hoffstaedter F, Nostro A, Yeo BTT, Eickhoff SB. 2020. Sex classification by resting state brain connectivity. *Cereb Cortex*. 30:824–835.

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.

Xu T, Falchier A, Sullivan EL, Linn G, Ramirez JSB, Ross D, Feczko E, Opitz A, Bagley J, Sturgeon D, et al. 2018. Delineating the macroscale areal organization of the macaque cortex in vivo. *Cell Rep*. 23:429–441.

Yeo BT, 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, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, et al. 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol*. 106:1125–1165.

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, Krienen FM, Eickhoff SB, Yaakub SN, Fox PT, Buckner RL, Asplund CL, Chee MWL. 2015. Functional specialization and flexibility in human association cortex. *Cereb Cortex*. 25:3654–3672.

Zuo X-N, Xing X-X. 2014. Test-retest reliabilities of resting-state FMRI measurements in human brain functional connectomics: a systems neuroscience perspective. *Neurosci Biobehav Rev*. 45:100–118.

Zuo X-N, Ehmke R, Mennes M, Imperati D, Castellanos FX, Sporns O, Milham MP. 2012. Network centrality in the human functional connectome. *Cereb Cortex*. 22:1862–1875.

Zuo X-N, Xu T, Milham MP. 2019. Harnessing reliability for neuroscience research. *Nat Hum Behav*. 3:768–771.