Estimates of segregation and overlap of functional connectivity networks in the human cerebral cortex

B.T. Thomas Yeo a,e,⁎, Fenna M. Krienen b,c, Michael W.L. Chee a, Randy L. Buckner b,c,d

a Center for Cognitive Neuroscience, Duke-NUS Graduate Medical School, Singapore
b Department of Psychology, Center for Brain Science, Harvard University, Cambridge, USA
c Department of Psychiatry, Massachusetts General Hospital, Charlestown, USA
d Department of Electrical and Computer Engineering, National University of Singapore, Singapore

Abstract

The organization of the human cerebral cortex has recently been explored using techniques for parcellating the cortex into distinct functionally coupled networks. The divergent and convergent nature of cortico-cortical anatomic connections suggests the need to consider the possibility of regions belonging to multiple networks and hierarchies among networks. Here we applied the Latent Dirichlet Allocation (LDA) model and spatial independent component analysis (ICA) to solve for functionally coupled cerebral networks without assuming that cortical regions belong to a single network. Data analyzed included 1000 subjects from the Brain Genomics Superstruct Project (GSP) and 12 high quality individual subjects from the Human Connectome Project (HCP). The organization of the cerebral cortex was similar regardless of whether a winner-take-all approach or the more relaxed constraints of LDA (or ICA) were imposed. This suggests that large-scale networks may function as partially isolated modules. Several notable interactions among networks were uncovered by the LDA analysis. Many association regions belong to at least two networks, while somatomotor and early visual cortices are especially isolated. As examples of interaction, the precuneus, lateral temporal cortex, medial prefrontal cortex and posterior parietal cortex participate in multiple paralimbic networks that together comprise subsystems of the default network. In addition, regions at or near the frontal eye field and human lateral intraparietal area homologue participate in multiple hierarchically organized networks. These observations were replicated in both datasets and could be detected (and replicated) in individual subjects from the HCP.

© 2013 Elsevier Inc. All rights reserved.

Introduction

Distributed neocortical brain areas form large-scale networks that exhibit complex patterns of divergent and convergent connectivity (e.g., Felleman and Van Essen, 1991; Goldman-Rakic, 1988; Jones and Powell, 1970; Mesulam, 1981; Pandya and Kupers, 1969; Ungerleider and Desimone, 1986). A major challenge in systems neuroscience is to make sense of these connectivity patterns to infer functional organization. In the visual system, connectivity patterns suggest a separation of processing into largely parallel, but interacting, hierarchical pathways (Felleman and Van Essen, 1991; Ungerleider and Desimone, 1986). In contrast, the association cortex comprises networks of widely distributed and densely interconnected areas without rigid hierarchical organization (Goldman-Rakic, 1988; Solomon and Goldman-Rakic, 1988; but see Badre and D’Esposito, 2009).

Resting-state functional connectivity MRI (rs-fcMRI) provides a powerful, albeit indirect, approach to make inferences about human cortical organization (Biswal et al., 1995). Despite its limitations (Buckner et al., 2013), we and others have used functional connectivity to estimate cortical network patterns (e.g., Bellec et al., 2010; Damoiseaux et al., 2006; He et al., 2009; Margulies et al., 2007; Power et al., 2011; Smith et al., 2009; van den Heuvel et al., 2009; Yeo et al., 2011).

The majority of functional connectivity studies have focused on dissociating functionally distinct networks or modules (Beckmann et al., 2005; Calhoun et al., 2008; Craddock et al., 2012; Damoiseaux et al., 2006; De Luca et al., 2006; Dosenbach et al., 2007; Doucet et al., 2011; Fox et al., 2006; Greicius et al., 2003; Margulies et al., 2007; Rubinov and Sporns, 2011; Salvador et al., 2005; Seeley et al., 2007; Smith et al., 2009; van den Heuvel et al., 2009; Varoguas et al., 2011). Fewer studies have examined the relationships among different functional networks (Sepulcre et al., 2012a; Sporns, 2013). For example, Fox et al. (2005) and Fransson (2005) have investigated the antagonistic relationship between the default and task-positive networks. Others (Doucet et al., 2011; Lee et al., 2012; Meunier et al., 2009) have investigated the (spatial) hierarchical relationship across functional networks. We previously employed a mixture model that relied on a winner-takes-all assumption to map network topology in the human cerebral cortex (Yeo et al., 2011). Each brain region was assigned to a single, best-fit network allowing us to derive connectivity maps that emphasize the...
interdigitation of parallel, distributed association networks. The key features of this parallel organization are that (1) each association network consists of strongly coupled brain regions spanning frontal, parietal, temporal, and cingulate cortices, and (2) the components of multiple networks are spatially adjacent (Yeo et al., 2011; also see Vincent et al., 2008; Power et al., 2011). However, it is unlikely that the brain is simply parcellated into a discrete number of nonoverlapping networks (Mesulam, 1998). Interactions across networks, as well as the existence of ‘convergence zones’ of regions that participate in multiple networks, are likely important features of brain organization (Beckmann et al., 2005; Bullmore and Sporns, 2009; Fornito et al., 2012; Jones and Powell, 1970; Mesulam, 1998; Pandya and Kuypers, 1969; Power et al., 2013; Sepulcre et al., 2012b; Spreng et al., 2010). Relevant to this point, we have observed variability in the goodness of fit of certain regions to their winner-takes-all network (Figs. 8 and 10 of Yeo et al., 2011), consistent with the notion that certain brain regions might participate in multiple networks (Andreas-Hanna et al., 2010; Beckmann et al., 2005; Leech et al., 2011; Rubinov et al., 2011; Sporns et al., 2007).

Here, we address the possibility of multiple network membership by applying latent Dirichlet allocation (LDA; Blei et al., 2003) and spatial Independent Component Analysis (ICA; Calhoun et al., 2001; Beckmann and Smith, 2004) to examine the topography of overlapping networks. This is an important consideration because network topography may change substantially from our original estimates (Yeo et al., 2011) if constraints are relaxed to permit overlapping networks. Conversely, unbiased estimation of network topography may broadly confirm previous estimates and allow us to investigate the interactions and overlaps among networks.

Materials and methods

Overview

We applied the LDA model to resting-state data from 1000 healthy young adults from the Brain Genomics Superstruct Project (GSP), as well as to 12 high quality, high-resolution individual subject datasets from the Human Connectome Project (HCP; Van Essen et al., 2013). The large sample size in GSP and the multiple sessions of individual HCP subjects permitted us to quantify patterns of cortico-cortical coupling that reveal insights into interactions within and across functional networks. Analyses proceeded in four stages. First, we applied the mixture model (Yeo et al., 2011) and LDA model (Blei et al., 2003) to both the GSP and HCP group datasets, in order to examine how cortical network organization changes as regions are permitted to participate in multiple networks (Fig. 1). For this analysis, the GSP and HCP datasets were used to provide independent replication samples. Next, we further analyzed several cortical regions participating in multiple sub-networks (Figs. 2 to 4). We then exploited the high quality, multi-session HCP data to determine if network organization can be estimated and replicated in individual subjects (Figs. 5 and 6). This increased the confidence that the discovered network organization was not merely a consequence of averaging across subjects. Additional control analyses confirmed similar network organization regardless of whether global signal regression was performed during preprocessing (Supplemental Fig. 7) and across degenerate (i.e., not highest likelihood) network estimates (Figs. 7 and 8).

Datasets

The GSP subjects were between ages 18–35 (mean age = 21.3; 42.7% male). Participants underwent one or two runs of eyes open rest (EOR). Analyses of the GSP data have been published previously (e.g., Buckner et al., 2011; Choi et al., 2012; Yeo et al., 2011). The HCP subjects were between ages 26–35 (mean age estimate = 30.9; 16.7% male). HCP provides aggregated data concerning age, hence mean age can only be estimated. HCP participants underwent two runs of passive fixation (FIX) in each of two separate sessions, for a total of four runs (~24 h interval between sessions).

GSP MRI data acquisition and preprocessing

Data were acquired on 3 T Tim Trio scanners (Siemens, Erlangen, Germany) using a 12-channel phased-array head coil. Functional data consisted of gradient-echo planar images (EPI) sensitive to blood oxygenation level-dependent (BOLD) contrast. Parameters for the resting data were: repetition time (TR) = 3000 ms, echo time (TE) = 30 ms, flip angle (FA) = 95°, 3 × 3 × 3 mm voxels, field of view (FOV) = 216, and 47 axial slices collected with interleaved acquisition. Slices were oriented along the anterior commissure–posterior commissure plane. Functional runs lasted 6.2 min (124 time points). Structural data included a multiecho T1-weighted magnetization-prepared gradient-echo (MP-RAGE) image (van der Kouwe et al., 2008).

fMRI processing steps included 1) discarding the first four frames of each run, 2) correcting for slice acquisition-dependent time shifts in each volume with SPM2 (Wellcome Department of Cognitive Neurology, London, UK), and 3) correcting for head motion using rigid body translation and rotation parameters (FSL; Jenkinson et al., 2002; Smith et al., 2004). This was followed by standard functional connectivity preprocessing (Fox et al., 2005; Van Dijk et al., 2010; Vincent et al., 2006). Linear trends over each run were removed and a low-pass temporal filter retained frequencies below 0.08 Hz. Spurious variance was removed using linear regression with terms for head motion, whole brain signal, ventricle signal, white matter signal and their derivatives.

Individual participants’ T1 scans were reconstructed into surface representations using FreeSurfer (http://surfer.nmr.mgh.harvard.edu; Fischl, 2012). Functional data were registered to structural images using FreeSurfer’s FmriBinarize package (Greve and Fischl, 2009; http://surfer.nmr.mgh.harvard.edu/fswiki/Fast). The structural preprocessing and structural–functional data alignment steps were described in Yeo et al. (2011). Functional data were projected onto the FreeSurfer surface space (2 mm mesh), smoothed on the surface using a 6 mm full-width half-maximum kernel, and were then downsampled to a 4 mm mesh.

HCP MRI data acquisition and preprocessing

HCP data were part of the HCP initial October 2012 public data release (http://www.humanconnectome.org/data). Data were acquired on a 3 T Skyra scanner (Siemens, Erlangen, Germany) using a standard 32-channel head coil. The scanner has a customized SGE2 gradient insert and a customized body transmitter coil with 56 cm bore size. The HCP Skyra has the standard set of Siemens shim coils (up to 2nd order). Functional data consisted of gradient-echo EPI sensitive to BOLD contrast. Parameters for the resting data were: TR = 720 ms, TE = 33.1 ms, FA = 52°, 2 × 2 × 2 mm voxels, FOV = 208 × 180 mm, and 72 oblique axial slices alternated between phase encoding in a right to left direction in one run and phase encoding in a left to right direction in the other run (Feinberg et al., 2010; Moeller et al., 2010; Setsompop et al., 2012; Xu et al., 2012). Each functional run lasted 14.55 min (1200 time points). Structural data included a T1-weighted MP-RAGE image. Parameters for the structural scan were as follows: TR = 2400 ms, TI = 1000 ms, TE = 2.14 ms, FA = 8°, 0.7 × 0.7 × 0.7 mm voxels and FOV = 224 × 224 × 224 mm. More details of the acquisition strategy can be found in Van Essen et al. (2012).

We utilized the fMRI preprocessed data released by the HCP (Glasser et al., 2013). fMRI processing steps included 1) gradient distortion correction (Jovicich et al., 2006, 2) motion correction, 3) distortion correction, 4) registration to the T1 scan (Greve and Fischl, 2009), 5) spline resampling to FSL MNI152 2 mm space using FSL FNIRT (Jenkinson et al., 2002; Smith et al., 2006), and 6) intensity normalization to mean of 10,000 and bias field correction. This was followed by standard functional connectivity preprocessing as in the GSP dataset. The preprocessed...
fcMRI data were projected from FSL MNI152 space onto the FreeSurfer surface space (1 mm mesh), smoothed using a 6-mm full-width half-maximum kernel and downsampled to a 4 mm mesh. The nonlinear mapping between FSL MNI152 volumetric space and FreeSurfer surface space is detailed in Buckner et al. (2011).

Clustering

The method of clustering cerebral cortical data has been previously described (Yeo et al., 2011). Briefly, for each subject, the Pearson's product-moment correlation was computed between each surface vertex \((N = 18,715)\) and 1175 uniformly distributed cortical regions of interest (ROIs). The “connectivity profile” of each surface vertex was its functional coupling to these ROIs. Each participant’s \(1175 \times 18,715\) matrix of correlations was binarized to retain the top 10% of correlations before summing across subjects to obtain an overall group estimate \(P\). Therefore, the \(i\)th row and \(j\)th column of the matrix \(P\) was the number of subjects whose correlations between the \(i\)th vertex and \(j\)th ROI are within the top 10% of correlations (within each individual subject). In other words, each matrix component took on integer values from 0 to 1000 in the GSP dataset. The connectivity profiles were clustered using a mixture of von Mises–Fisher distributions (Lashkari et al., 2010; Yeo et al., 2011). We repeated the clustering algorithm 1000 times with different random initializations and selected the estimate with best model likelihood. For more details, we refer readers to Lashkari et al. (2010) and Yeo et al. (2011). Because our previous analyses (Yeo et al., 2011) identified solutions with 7 and 17 network clusters to be particularly stable, we adopted these for the present study.

**Latent Dirichlet Allocation (LDA)**

LDA was first introduced in the text mining literature (Blei et al., 2003). The application of LDA to estimate overlapping modules (networks) in graphs has been previously proposed (Zhang et al., 2007). Here, we employed LDA to estimate cortical networks from resting-state fMRI data.

Like spatial independent component analysis (ICA; Beckmann and Smith, 2004), LDA permits a brain region to belong to multiple networks. Both ICA (Beckmann and Smith, 2004) and LDA (Hoffman et al., 2010) seek to factorize a matrix \(M\) into a product of two matrices \(W\) and \(H\). Because there are an infinite number of ways to factorize a matrix, additional constraints are required.

In the case of spatial ICA applied to fMRI, the spatial weights of the estimated networks (components) are constrained to be independent (Beckmann et al., 2005; Calhoun et al., 2001), but can take on negative values. LDA is closely related to non-negative matrix factorization (Hoffman et al., 2010). When applied to resting-state fMRI data in this paper, LDA constrains the spatial weights of the estimated networks to be non-negative (in addition to other constraints). We refer interested readers to Blei et al. (2003) for the probabilistic (and more well-known) interpretation of LDA.

In this work, we used LDA to factorize the \(1175 \times 18,715\) matrix \(P\) (see the previous “Clustering” section) into the product of two matrices \(W (1175 \times K)\) by \(H (K \times 18,715)\), where \(K\) is the number of networks. The \(k\)-th row of the matrix \(H\) sums to one over all \(18,715\) vertices of the cerebral cortex. We can interpret the \(k\)-th row as the probability that the \(k\)-th functional connectivity network includes a particular surface vertex \(Pr(\text{vertex} | \text{k-th network})\).

To compare the clustering and LDA results, we consider 7 and 17 networks in this paper. The LDA model is estimated with 100 random initializations. The estimate with the best likelihood bound is selected. The LDA code is publicly available.1

---

1 http://www.cs.princeton.edu/~blei/lda-c/
2 Using a different measure (e.g., KL-divergence) achieves the same results.

Spatial independent component analysis (spatial ICA)

We applied spatial ICA (FSL melodic 3.10; Beckmann and Smith, 2004) to the surface projected fcMRI-preprocessed data. We estimated 20 ICA components because of precedence (e.g., Smith et al., 2009) and because “20” is close to the number of networks (“17”) we sought in the clustering and LDA models. We also experimented with 17 independent components, as well as applying ICA directly to raw volumetric fMRI data before projecting the spatial components to the surface. The results were similar across these different experiments and so for conciseness, we will focus on the 20 ICA components estimated from surface-projected fcMRI-preprocessed data.

Matching networks between datasets and methods

Since the ordering of the networks (or components) estimated using different methods and datasets is arbitrary, we used the Hungarian matching algorithm (Kuhn, 1955) to find the correspondences between networks estimated with different methods and datasets. Essentially, the networks were relabeled so as to maximize the spatial agreement between corresponding networks.

In particular, HCP clustering estimates were reordered to match the GSP clustering estimates by maximizing the number of vertices belonging to corresponding networks across the datasets. Similarly, the HCP LDA estimates were reordered to match the GSP LDA estimates by maximizing the correlation between \(Pr(\text{vertex} | \text{network})\) of corresponding networks across datasets.2

To match clustering estimates with LDA estimates, we thresholded the LDA estimates to obtain winner-takes-all parcellations. The clustering estimates were then reordered to match the winner-takes-all LDA estimates by maximizing the number of vertices belonging to corresponding (winner-takes-all) networks across methods.

Finally, we found 17 of the 20 ICA components that best correspond to the 17 LDA estimates by maximizing the correlation between \(Pr(\text{vertex} | \text{LDA network})\) and the un-thresholded, Z-score ICA maps of corresponding networks across methods.

Visualization

Clustering, LDA and ICA estimates were transformed from FreeSurfer surface space (Dale et al., 1999; Fischl et al., 1999a,b) to the PALS-B12 surface space for visualization using Caret (Van Essen, 2005). Network colors match Yeo et al. (2011). Lines representing clustering boundaries were smoothed to remove high spatial frequency jaggedness resulting from mapping the data to PALS space.

Comparison of network estimates in GSP and HCP datasets

The 7-network clustering and LDA analyses of the GSP and HCP datasets are shown in Fig. 1. Since each LDA network comprises a probability distribution over vertices \(Pr(\text{vertex} | \text{network})\), we can visualize each distribution as a map over the cortex. Because LDA is a Bayesian model, there is no obvious threshold for displaying the probability maps. Given that there are 18,715 vertices and the probability distribution sums to one over all the vertices, any probability larger than 1/18,715 is above expectation. Here, we chose a relatively stringent threshold of 1e−4 (roughly two times 1/18,715). The LDA estimates in Fig. 1 were overlaid on the boundaries of the respective 7-network clustering estimates in the GSP and HCP datasets. The boundaries allowed the comparison between the clustering and LDA estimates.

The 17-network clustering and LDA analyses of the GSP and HCP datasets are shown in Figs. 3 and 4. Because of the small number of
subjects in the HCP dataset, we will focus our interpretation on the similarities, rather than the differences, between the HCP and GSP estimates.

Supplemental Figs. 1 and 2 juxtapose 7 of the 17 corresponding LDA networks and ICA components in the GSP and HCP datasets respectively. To foreshadow the results, ICA and LDA provide very similar network estimates, except for negative spatial weights in the ICA estimates. Both ICA and LDA decompose a signal into a linear combination of component signals (Lee and Seung, 1999). For LDA, only additive combinations are allowed. This is in contrast with ICA, which may allow component signals to cancel each other out via subtractions (Lee and Seung, 1999). It may be the case that the negative spatial weights in ICA are functionally meaningful, but assessment is difficult since ground-truth network estimates are not available. Consequently, we chose to focus on the LDA results.3

Robustness to number of networks

As discussed in our previous paper (Yeo et al., 2011), the focus on 7-network and 17-network solutions should not be taken to imply that meaningful properties are absent in alternative parcellation schemes. By focusing on both a relatively coarse solution (7 networks) and a fine-resolution solution (17 networks), we hoped to capture broad aspects of the solution space that were consistent within these relative extremes. However, we observed stable results for other network parcellations as well (e.g. 10 & 12 networks, see Fig. 6 of Yeo et al., 2011). To ensure our present results are robust to the choice of 7 and 17 networks, we estimated 4, 6, 8, 10, 12, 16, 18 and 30 networks by applying clustering and LDA to the GSP dataset (Supplemental Fig. 3).

Quantifying overlap between networks

To assess the overlap between networks, we computed \( \Pr(\text{network} | \text{vertex}) \) for each vertex by applying Bayes’ rule to \( \Pr(\text{vertex} | \text{network}) \) and assuming each network is equally likely, i.e., \( \Pr(\text{network}) = 1/K \). Since \( \Pr(\text{network} | \text{vertex}) \) sums to 1 over the estimated K networks, any probability larger than 1/K is above expectation. For each vertex, we consider the number of networks with \( \Pr(\text{network} | \text{vertex}) \) above 1/K as a rough measure of the number of networks the vertex is participating in. Fig. 2A shows the map of the number of networks each vertex participates in. For this analysis we focus on the most stable solution — 7-network LDA results for the GSP dataset.

To quantify the distribution of multiple-network participation across the cerebral cortex, we computed the fraction of vertices participating in more than one LDA network for each network of the 7-network clustering estimates (Fig. 2B). Vertices within 10 mm of clustering boundaries were excluded from this analysis because they might reflect uncertainties in the network estimates. We also computed the fraction of vertices participating in more than one LDA network for each winner-takes-all LDA network (Supplemental Fig. 4). Vertices within 10 mm of the winner-takes-all LDA boundaries were excluded from this analysis.

To ensure the above analysis was robust to the choice of threshold for \( \Pr(\text{network} | \text{vertex}) \), we repeated the analysis using a more liberal threshold of 0.75/K and a more conservative threshold of 1.25/K (Supplemental Fig. 5). We also repeated the analysis with 6-network and 8-network LDA estimates to ensure our results were robust to the choice of number of networks (Supplemental Fig. 6).

As the results will show, the default and dorsal attention networks have the greatest proportion of regions participating in multiple networks. Therefore, we identified brain regions that (1) were at least 10 mm away from both 7-network clustering and winner-takes-all LDA boundaries, (2) participated in multiple networks from the 7-network estimate and either (3a) participated in multiple networks (from the 17-network LDA estimate) that overlapped significantly with the default network or (3b) participated in multiple networks (from the 17-network LDA estimate) that overlapped significantly with the dorsal attention network. We explored the participation of these regions in multiple networks from the 17-network LDA estimates (Figs. 3 and 4).

LDA estimates in individual subjects

To ensure the network organization we discovered was not an artifact of intersubject averaging, we applied the 7-network LDA model to individual sessions of each HCP subject. This also allowed us to evaluate the test–retest reliability of the LDA estimates. The LDA estimates of both sessions of subjects with the best (Fig. 5) and median (Fig. 6) test–retest reliability are juxtaposed.

Global signal regression

The regression of global signal during fcMRI preprocessing can introduce negative correlations between brain regions (Fox et al., 2009; Murphy et al., 2009). To ensure the network organization we discovered was robust to variations in fcMRI preprocessing, we randomly selected a subset of 100 subjects from the GSP dataset and processed them without global signal regression. We referred to this data as the non-GSR dataset. We then estimated 7 and 17 networks from this non-GSR dataset using both von Mises–Fisher clustering and LDA (Supplemental Fig. 7).

Degenerate high-likelihood clustering and LDA estimates

While we have focused on network estimates with the highest likelihood, there might be degenerate or alternate network estimates with high likelihood values that are topologically distinct from the best estimates (Good et al., 2010). Given that these degenerate estimates have likelihood values close to the best estimate, slight variations of the original models might result in these alternate network estimates becoming the best solutions. Consequently, we employed two different approaches to explore whether the network organization of the degenerate solutions agrees with that of the best estimate (cf. Rubinov and Sporns, 2011).

First, we assessed whether degenerate estimates exhibited parallel large-scale networks in association cortex by exploring the solution space of the von Mises–Fisher mixture model obtained from 1000 different random initializations. We computed the overlap agreement between each of the 1000 estimates and the best estimate in the GSP dataset (Fig. 7A). In addition, we computed an “exhaustive” set of 7-network clustering estimates that are representative of the 1000 random initializations (Fig. 7B). We determined this “exhaustive” set as follows. We first initialized the “exhaustive” set to consist only of the best clustering estimate (i.e., estimate with the highest likelihood). We then iterated the following procedure until convergence: At each iteration, among all clustering estimates with less than 80% overlap with the current “exhaustive” set of clustering estimates, we added the clustering estimate with the highest likelihood to the “exhaustive” set. The procedure therefore allowed us to select a subset of clustering solutions that were the most different from each other and spanned the solution space of the mixture model. The 80% overlap threshold was chosen because we found (empirically) that estimates with more than 80% overlap were extremely similar (visually).

Second, to assess whether degenerate estimates exhibited the property that many association regions participated in multiple networks,
we computed the number of networks each vertex participates in, averaged across the 100 random initializations of the 7-network and 17-network LDA (Fig. 8).

Results

Clustering estimates of GSP and HCP datasets

The 7-network clustering estimates of the GSP and HCP datasets were similar (top row of Fig. 1), with 78% of vertices identically labeled within the entire cerebral cortex. Five of the networks were highly similar across the two datasets, with overlaps between the somatomotor (blue) and visual (purple) networks at more than 90%. Notable differences included portions of the default (red) network in the GSP dataset classified as part of the limbic (cream) network in the HCP dataset. This may relate to differences in signal-to-noise around regions of high susceptibility (which severely influences estimates of the limbic network). The overlap of the 17-network clustering estimates (not shown) between the GSP and HCP datasets was 69%.

Fig. 1. Large-scale network organization is reliable across datasets (GSP and HCP) and across methods (clustering and LDA). Top row shows the 7-network clustering estimates of the GSP and HCP datasets in the left hemisphere. Overlap (fraction of vertices with same network labels) between GSP and HCP estimates was 78% across the entire cerebral cortex. Five (purple, blue, green, violet and orange) of the seven networks were virtually identical. Bottom rows show the 7-network LDA estimates of the GSP and HCP datasets overlaid on blue boundaries corresponding to the clustering estimates of respective datasets. All seven LDA networks were highly similar across the two group datasets. The mean of the Pearson correlation coefficients of the seven pairs of LDA network estimates was 0.93. The colored squares (center column) indicate correspondences between the LDA and clustering estimates. For example, the LDA estimate in the second row corresponds to the purple visual cluster. The percent overlap between the clustering and winner-takes-all LDA estimates in the GSP dataset was 82%, while the overlap in the HCP dataset was 88%.
The 7-network LDA estimates of the GSP and HCP datasets were similar (rows 2–8 of Fig. 1). The mean of the Pearson correlations between 17 corresponding pairs of networks were 0.72 in the GSP dataset and 0.62 in the HCP dataset. Given the similarities between ICA and LDA (except for the negative spatial weights in the ICA estimates), we will focus on the LDA estimates.

ICA estimates of GSP and HCP datasets

The LDA and ICA estimates were similar in the GSP (Supplemental Fig. 1) and HCP (Supplemental Fig. 2) datasets. The means of the Pearson correlations between 17 corresponding pairs of networks were 0.72 in the GSP dataset and 0.62 in the HCP dataset. The agreement between the 7-network clustering and LDA estimates can be visually appreciated by looking at the correspondences between the LDA estimates and clustering boundaries (rows 2–8 in Fig. 1).

The overlap between the 7-network clustering estimates and the winner-takes-all LDA estimates in the GSP dataset was 82%, while the overlap in the HCP dataset was 88%. The overlap between the 17-network clustering estimates and the winner-takes-all LDA estimates in the GSP dataset was 70%, while the overlap in the HCP dataset was 63%.

Robustness to number of networks

The focus on 7- and 17-network estimates should not be taken to imply that meaningful properties are absent in alternative parcellation schemes. We found similar agreement between the clustering and LDA estimates for 4, 6, 8, 10, 12, 16, 18 and 30 networks in the GSP dataset; the overlap agreement between the clustering and LDA estimates for the different number of networks.

The overlap in the HCP dataset was 88%. The overlap between the 17-network clustering estimates and the winner-takes-all LDA estimates in the GSP dataset was 70%, while the overlap in the HCP dataset was 63%.

A critical question is the degree to which the winner-takes-all and LDA approaches converge on the same basic network organization. The agreement between the 7-network clustering and LDA estimates can be visually appreciated by looking at the correspondences between the LDA estimates and clustering boundaries (rows 2–8 in Fig. 1).

The overlap between the 7-network clustering estimates and the winner-takes-all LDA estimates in the GSP dataset was 82%, while the overlap in the HCP dataset was 88%. The overlap between the 17-network clustering estimates and the winner-takes-all LDA estimates in the GSP dataset was 70%, while the overlap in the HCP dataset was 63%.

Robustness to number of networks

The focus on 7- and 17-network estimates should not be taken to imply that meaningful properties are absent in alternative parcellation schemes. We found similar agreement between the clustering and LDA estimates for 4, 6, 8, 10, 12, 16, 18 and 30 networks in the GSP dataset; the overlap agreement between the clustering and LDA estimates for the different number of networks.

The overlap in the HCP dataset was 88%. The overlap between the 17-network clustering estimates and the winner-takes-all LDA estimates in the GSP dataset was 70%, while the overlap in the HCP dataset was 63%.

A critical question is the degree to which the winner-takes-all and LDA approaches converge on the same basic network organization. The agreement between the 7-network clustering and LDA estimates can be visually appreciated by looking at the correspondences between the LDA estimates and clustering boundaries (rows 2–8 in Fig. 1).

The overlap between the 7-network clustering estimates and the winner-takes-all LDA estimates in the GSP dataset was 82%, while the overlap in the HCP dataset was 88%. The overlap between the 17-network clustering estimates and the winner-takes-all LDA estimates in the GSP dataset was 70%, while the overlap in the HCP dataset was 63%.

Robustness to number of networks

The focus on 7- and 17-network estimates should not be taken to imply that meaningful properties are absent in alternative parcellation schemes. We found similar agreement between the clustering and LDA estimates for 4, 6, 8, 10, 12, 16, 18 and 30 networks in the GSP dataset; the overlap agreement between the clustering and LDA estimates for the different number of networks.

The overlap in the HCP dataset was 88%. The overlap between the 17-network clustering estimates and the winner-takes-all LDA estimates in the GSP dataset was 70%, while the overlap in the HCP dataset was 63%.

A critical question is the degree to which the winner-takes-all and LDA approaches converge on the same basic network organization. The agreement between the 7-network clustering and LDA estimates can be visually appreciated by looking at the correspondences between the LDA estimates and clustering boundaries (rows 2–8 in Fig. 1).

The overlap between the 7-network clustering estimates and the winner-takes-all LDA estimates in the GSP dataset was 82%, while the overlap in the HCP dataset was 88%. The overlap between the 17-network clustering estimates and the winner-takes-all LDA estimates in the GSP dataset was 70%, while the overlap in the HCP dataset was 63%.
network and the limbic network had below average number of cortical vertices participating in multiple networks.

We obtained consistent results when we computed the fraction of vertices participating in multiple LDA networks for each winner-takes-all LDA network (Supplemental Fig. 4) or when we repeated the analyses with different thresholds (Supplemental Fig. 5) or different number of networks (Supplemental Fig. 6).

Across analyses (Fig. 2B and Supplemental Figs. 4 to 6), the default and dorsal attention networks had the largest proportions of brain vertices participating in multiple networks. To explore this phenomenon further, we identified brain regions in the left hemisphere that (1) were at least 10 mm away from both clustering and winner-takes-all LDA boundaries, (2) participated in multiple networks from the 7-network estimate and (3) participated in multiple networks (from the 17-network LDA estimate) that overlapped significantly with the default network (from the 7-network estimate). Four regions were identified (black asterisks in Fig. 2A): precuneus (PCUN), lateral temporal cortex (LTC), posterior parietal cortex (PPC) and medial prefrontal cortex (MPFC). Their MNI coordinates are reported in Table 1. It is worth pointing out that the PCUN region was surrounded by a band of regions participating in only one network, suggesting that its participation in multiple networks was not due to spatial blurring across network boundaries.

We also selected brain regions in the left hemisphere that (1) were at least 10 mm away from both clustering and winner-takes-all LDA boundaries, (2) participated in multiple networks from the 7-network estimate and (3) participated in multiple networks (from the 17-network LDA estimate) that overlapped significantly with the dorsal attention network (from the 7-network estimate). Two regions were identified (blue asterisks in Fig. 2A) and they are putative homologues of macaque areas LIP (lateral intraparietal) and FEF (frontal eye fields) based on a meta-analysis of fMRI literature (Yeo et al., 2011). Coordinates of these two regions are reported in Table 1.

We now explore the participation of the default regions in multiple paralimbic networks and the dorsal attention regions in multiple hierarchically organized networks in the 17-network LDA estimates.

Fig. 3. The precuneus (PCUN), lateral temporal cortex (LTC), posterior parietal cortex (PPC) and medial prefrontal cortex (MPFC) participate in multiple subnetworks. (A) 17-network clustering estimates of the GSP and HCP datasets. Only left hemisphere estimates are shown. The percent overlap between the clustering estimates in the two datasets was 69%. (B) Four of the 17-network LDA estimates overlapping traditional default network regions are shown. The LDA networks in the 1st, 2nd and 4th panels overlap with the yellow, red and dark-blue networks in (A) respectively. Panel 3 shows the LDA network that overlaps with the gold and cream networks in (A). The asterisks correspond to the default regions (PCUN, LTC, PPC and MPFC) defined in Table 1. An asterisk is colored black if \( \Pr(\text{vertex} | \text{network}) \geq 1 \times 10^{-4} \) at that brain location.
Overlapping components of multiple paralimbic networks

The 17-network clustering estimates in the GSP and HCP datasets are juxtaposed in Fig. 3A. Fig. 3B shows four paralimbic networks from the 17-network LDA estimates. The networks overlapped significantly with brain regions typically associated with the default network. The asterisks (Fig. 3B) correspond to the default regions defined in Table 1. An asterisk is colored black if $\Pr(\text{vertex} | \text{network}) > 1 \times 10^{-4}$ at that brain location. Therefore both the PCUN and PPC regions preferentially participated in paralimbic networks 1, 2 and 4; the MPFC region preferentially participated in paralimbic networks 1, 3 and 4; the LTC region preferentially participated in all four paralimbic networks.

FEF and LIP are involved in multiple hierarchically organized networks

The 17-network clustering estimates in the GSP and HCP datasets are juxtaposed in Fig. 4A. Fig. 4B shows four networks from the 17-network LDA estimates that overlap significantly with brain regions typically associated with the dorsal attention network. These networks are likely the homologue of the well-studied hierarchical visuomotor pathway in macaques (Maunsell and Van Essen, 1983). The asterisks (Fig. 4B) correspond to the dorsal attention regions defined in Table 1. An asterisk is colored black if $\Pr(\text{vertex} | \text{network}) > 1 \times 10^{-4}$. Therefore the LIP-homologue region preferentially participated in all four networks, while the putative FEF region preferentially participated in networks 2, 3 and 4. The visual hierarchy from early visual cortex to superior parietal cortex and frontal eye fields was noted in Yeo et al. (2011) via a series of seed-based functional connectivity analyses. As this example illustrates, LDA detected overlap among functional networks comprising early retinotopic visual areas and other components of this sensory-motor processing stream that were lost when making a winner-take-all assumption.

Cortical network structure can be detected in individual subjects

Many research goals require analysis of imaging data at the individual subject level. The HCP dataset provides high quality datasets that are obtained twice in participants over separate days. To ensure the network organization we discovered was not an artifact of averaging across subjects, we analyzed these data to determine whether LDA networks could be detected within individual subjects and whether they would show reliability.

Pairwise correlation of the 7-network LDA estimates between the two sessions of each subject was 0.60 on average. The best subject had a mean pairwise correlation of 0.75. The median subject had a mean pairwise correlation of 0.61. Nine of the twelve subjects had average correlations of at least 0.56, suggesting moderate reliability.

The 7-network LDA estimates of the best and median subjects are shown in Figs. 5 and 6. The estimated networks were well
replicated across the two sessions even for the median subject. This is consistent with previous assessments of test-retest reliability of resting-state networks (Shehzad et al., 2009; Zuo et al., 2010).

In both subjects (Figs. 5 and 6), the visual (row 1), somatomotor (row 2) and paralimbic (rows 6 and 7) networks were well replicated. However, the dorsal attention (green), ventral attention (violet) and frontoparietal control (orange) networks tended to inter-mix across sessions (rows 3 to 5 of Figs. 5 and 6). This is not surprising given that the regions from these networks are correlated (Fox et al., 2006; Yeo et al., 2011). For example, dorsal attention regions are known to correlate with ventral attention regions (Fox et al., 2006; Yeo et al., 2011).

**Choice of thresholding LDA estimates**

We should emphasize that the choice of thresholding the LDA networks at $10^{-4}$ was not critical for our analyses. While thresholding $P_r(\text{vertex} \mid \text{network})$ networks was necessary to visualize the results, statistics of agreement among different data-driven approaches and datasets (Fig. 1, Supplemental Figs. 1 and 2) were computed independent of the threshold. Similarly, statistics for test-retest reliability (Figs. 5 and 6) were computed independent of the threshold. Finally, given the relatively stringent threshold of $10^{-4}$ (almost twice the minimal threshold of $1/18,715$), using a lower threshold would simply confirm our broad point that different cortical regions identified in Table 1 were involved in multiple subnetworks (Figs. 3 and 4).
Cortical network structure is invariant to global signal regression

To ensure the network organization we discovered was not an artifact of global signal regression (GSR) during fMRI preprocessing, we estimated 7 and 17 networks in the non-GSR dataset consisting of 100 GSP subjects processed without global signal regression. The 7-network and 17-network clustering estimates from the non-GSR dataset are shown in Supplemental Fig. 7. We found excellent agreement between the network estimates from the non-GSR and full GSP datasets. In particular, the overlap of the 7-network clustering estimates was 95%, while the overlap of the 17-network clustering estimates was 91% across the two datasets. The mean pairwise correlation of the 7-network LDA estimates was 0.99, while the mean pairwise correlation of the 17-network LDA estimates was 0.85 across the two datasets.

Cortical network structure is replicated in degenerate network estimates

Finally, we assessed whether degenerate network estimates reaffirmed the network organization we discovered. Fig. 7A plots the overlap between the best 7-network clustering estimate and estimates derived from the 1000 random initializations using the GSP dataset. The clustering estimates were ordered from the highest likelihood (best estimate) to the smallest likelihood. Out of 1000 random initializations, 415

Fig. 6. Similar to Fig. 5, LDA estimates of the HCP subject with the median test–retest reliability is displayed. Format is identical to Fig. 5. The mean of the Pearson correlation coefficients of the seven pairs of network estimates was 0.61. The dotted lines indicate correspondences between the LDA networks across the two sessions. By visual inspection, the third, fourth and fifth networks in sessions 1 and 2 were inter-mixed across the sessions.
resulted in solutions with overlap of over 90% with the best solution, suggesting the best estimate corresponds to a robust local optimum.

Fig. 7B shows the “exhaustive” set of 7-network clustering estimates. The eight clustering estimates were selected to be different from each other and spanned the solution space of the mixture model. Even though the seven “degenerate” clustering estimates do not have the highest likelihoods, they still explained interesting variance in the data. For example, three degenerate solutions fractionate the somatomotor cortex into dorsal–ventral portions, just like in the best 17-network estimate (Fig. 3A).

Importantly, while the details of the clustering estimates varied, the broad property that early sensory and late motor cortices appeared to participate in preferentially local networks was also found across all degenerate solutions. Furthermore, (1) each association network spanned multiple lobes and (2) the components of multiple networks were spatially adjacent. Therefore, the broad property that the human association cortex consisted of multiple, parallel networks was found across all degenerate solutions.

Similarly, we found that across 100 random initializations of the 7-network and 17-network LDA (Fig. 8), many association regions appeared to participate in multiple networks, while large portions of early sensory and late motor cortices participated in single networks.

Discussion

Complex behaviors are subserved by distributed networks of specialized brain areas (Distler et al., 1993; Mesulam, 1998; Posner et al., 1988; Shadlen and Newsome, 2001). In this work, we show that the human association cortex consists of multiple, interdigitated distributed networks in contrast to early sensory and late motor cortices that participate in preferentially local networks (Fig. 1). Many association regions appear to participate in multiple networks, while large portions of early sensory and late motor cortices participate in single networks (Fig. 2; Supplemental Fig. 4). We demonstrate examples of overlapping paralimbic networks (Fig. 3) and overlapping hierarchically organized networks (Fig. 4). This architecture was detected and replicated in individual subjects (Figs. 5 and 6). Additional control analyses confirmed similar network organization regardless of global signal regression (Supplemental Fig. 7) and across degenerate network estimates.
participate in preferentially local networks (Fig. 1; also see Sepulcre et al., 2010). In contrast, the association cortex comprises interdigitated, partially overlapping large-scale circuits (Figs. 1 and 2; Supplemental Figs. 1 and 2). The key features of this organization are that (1) each association network consists of strongly coupled brain regions spanning frontal, parietal, temporal, and cingulate cortices, and (2) the components of multiple networks exist adjacent to each other and are partially overlapping.

This organization is replicable across datasets (GSP and HCP) and techniques (clustering, LDA and ICA; Fig. 1 and Supplemental Figs. 1 and 2). Although details may differ, we can replicate this general organization across different number of estimated networks (Supplemental Fig. 3), with and without global signal regression (Supplemental Fig. 7) and across degenerate network estimates (Fig. 7). Finally, the organization can be detected in individual subjects and replicated across sessions in the HCP dataset (Figs. 5 and 6), indicating that this parallel architecture is not an artifact of group averaging.

Given the convergence in cortical network organization across datasets and methods, it is surprising that this parallel organization has not been emphasized in previous ICA analyses. We suspect one possible reason is that the winner-takes-all approach allows us to juxtapose the spatial organization of different networks in one single map (top row of Fig. 1), as opposed to more sophisticated approaches allowing for multiple memberships (bottom rows of Fig. 1). The parallel architecture is easily appreciated in the top row of Fig. 1, but is less clear in the bottom rows, even though the estimated networks are extremely similar (88% overlap in the HCP dataset). Furthermore, most previous analyses (e.g., Beckmann et al., 2005; Damoiseaux et al., 2006; De Luca et al., 2006; Smith et al., 2009) visualized the estimated networks (components) using brain slices, making it difficult to appreciate complex topographic relationships.

### Estimates of segregation and overlap of functional connectivity networks

In general, network overlaps occur throughout association cortex (Fig. 2a). These overlaps do not exclusively co-localize with the presence of a border (Fig. 2b; Table 1; Supplemental Fig. 4). We exclude these boundary regions in our analyses (Fig. 2b; Table 1; Supplemental Fig. 4) because they may reflect uncertainties in the network estimations (c.f. Power et al., 2013). The default and dorsal attention networks have the largest proportions of brain vertices participating in multiple networks (Fig. 2b; Supplemental Fig. 4). These results can be replicated with different number of networks (Supplemental Fig. 5), different thresholds (Supplemental Fig. 6) and across degenerate network estimates (Fig. 8).

We will focus on the dorsal attention and default networks in the following sections. However, we should emphasize that regions in other association networks also participate in multiple networks. Our emphasis on the default and dorsal attention networks may simply be due to their large sizes and our exclusion of regions close to network boundaries. This necessarily precludes investigation of the smaller regions of any network, though examples of regions possessing multiple network membership can be found in all networks at the resolution examined.

For example, there are regions in the posterior cingulate sulcus portion of the ventral attention network (at or near area 5 Ci; Scheperjans et al., 2008a, 2008b) at least 10 mm away from any network boundaries, which also participate in the dorsal attention network (rows 4 and 5 of Fig. 1). This is consistent with the correlations (Fox et al., 2006; Yeo et al., 2011) and close functional relationships (Corbetta and Shulman, 2002; Corbetta et al., 2008) between the two networks. Corbetta et al. (2008), as part of a review, noted that a region of right prefrontal cortex overlapped the dorsal and ventral attention systems, providing a candidate region for allowing interactions between the otherwise distinct systems (also see Asplund et al., 2010). The present results suggest the cingulate is also a potential point of interaction with the caveat that, for all such observations, it is difficult to know whether

### Association cortex comprises interdigitated, segregated large-scale circuits

Our findings suggest that the early sensory and late motor cortices participate in preferentially local networks (Fig. 1; also see Sepulcre et al., 2010). The plots show the number of LDA networks each cortical region participates in for the 7-network and 17-network LDA estimates in the GSP dataset, averaged across 100 random initializations. Many association regions participated in multiple networks, while large portions of early sensory and late motor cortices participated in single networks.

(Figs. 7 and 8). The overlaps among networks may provide clues to the network interactions that support human cognition.

### Table 1

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Locations of default and dorsal attention network regions that participate in multiple networks.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brain region</td>
<td>MNI coordinates</td>
</tr>
<tr>
<td>----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Precuneus (PCUN)</td>
<td>−6, −57, 36</td>
</tr>
<tr>
<td>Lateral temporal cortex (LTC)</td>
<td>−58, −11, −15</td>
</tr>
<tr>
<td>Posterior parietal cortex (PPC)</td>
<td>−43, −67, 39</td>
</tr>
<tr>
<td>Medial prefrontal cortex (MPFC)</td>
<td>−11, 46, −3</td>
</tr>
<tr>
<td>Lateral intraparietal (LIP) area</td>
<td>−28, −61, 80</td>
</tr>
<tr>
<td>Frontal eye fields (FEF)</td>
<td>−26, −6, 48</td>
</tr>
</tbody>
</table>

Fig. 8. Pattern of network segregation and overlap replicated across degenerate LDA estimates. The plots show the number of LDA networks each cortical region participates in for the 7-network and 17-network LDA estimates in the GSP dataset, averaged across 100 random initializations. Many association regions participated in multiple networks, while large portions of early sensory and late motor cortices participated in single networks.
there is true spatial convergence or distinct modules are present below our level of resolution. With this caveat in mind, we now focus our attention on two examples of overlapping networks.

Overlapping parallel paralimbic networks

Evidence from both human imaging and non-human primate anatomical tract tracing suggest the existence of a network of brain regions, known as the default network, involved in internal mentation (for reviews, see Buckner et al., 2008; Binder et al., 2009). These regions include the medial temporal lobe, posterior medial cortex (precuneus and posterior cingulate cortex), medial prefrontal cortex, inferior parietal cortex and lateral temporal cortex.

Recent evidence (Andrews-Hanna et al., 2010; Laird et al., 2009; Leech et al., 2011; Yeo et al., 2011) suggests functional heterogeneity within the default network. In particular, Andrews-Hanna et al. (2010) suggested that the posterior cingulate cortex (PCC) and the (anterior) medial prefrontal cortex (MPFC) form the core of two overlapping default sub-networks. Our current analysis extends Andrews-Hanna’s results, suggesting that the precuneus (PCUN), medial prefrontal cortex (MPFC), lateral temporal cortex (LTC) and posterior parietal cortex (PPC) participate in multiple paralimbic networks that overlap traditional default network regions (Fig. 3). In particular, our analysis highlights the LTC, which appears to participate in all four paralimbic networks.

Previous results have suggested functional heterogeneity within the posterior medial cortex (Fornito et al., 2012; Laird et al., 2009; Leech et al., 2011; Margulies et al., 2009). Our results also suggest the presence of such heterogeneity. For example, the PCC appears to preferentially participate in paralimbic networks 1 and 4 (Fig. 3), while PCUN (asterisk in Fig. 3) preferentially participates in paralimbic networks 1, 2 and 4.

While we have focused on its involvement in paralimbic networks (Fig. 3), the PCUN region may also participate in a network that prominently includes the specific rostral regions of prefrontal cortex (row 5 in Supplemental Figs. 1 and 2), although the involvement is clearer in the LDA estimates than in the ICA estimates. Since the rostral prefrontal cortex is implicated in tasks requiring complex rule application (e.g., Badre and D’Esposito, 2007), this network might be involved in higher-order cognitive control. This may be consistent with heterogeneity of the precuneus previously reported, such as “Cognitive/Associative Central Precuneus” in Margulies et al. (2009) and “Fronto-parietal Network” in Fornito et al. (2012).

Overlapping hierarchical networks in a canonical sensory-motor pathway

While the early sensory and late motor cortices mostly participate in single networks (Fig. 2b; Supplemental Fig. 4), there are exceptions. For example, overlaps between the sensory and motor networks with the dorsal attention network in association cortex (Fig. 4) may be expected from the non-human primate connectivity anatomical studies (Fellman and Van Essen, 1991) and may be involved in the hierarchical flow of information among sensory and motor areas.

The canonical sensory-motor pathway, including retinotopic visual cortex, the MT+ complex, parietal area LIP and the FEF, has been well characterized in the monkey (Andersen et al., 1990; Colby and Goldberg, 1999; Gold and Shadlen, 2007; Maunsell and Van Essen, 1983; Shadlen and Newsome, 2001). The idea is that visual information propagates from early visual areas to MT+, which constrains decision processes arising from interactions with LIP and FEF in the upper stages of the hierarchy. The anatomical connectivity among the areas within this pathway has been extensively explored in nonhuman primates (Fellman and Van Essen, 1991; Maunsell and Van Essen, 1983). In the human literature, this pathway has been studied both in relation to spatially directed movements and also in relation to spatial attention, with components of the pathway sometimes referred to as the dorsal attention network (Corbetta and Shulman, 2002).

This visual hierarchy was tested in Yeo et al. (2011) via a series of seed-based functional connectivity (also see Sepulcre et al., 2012b). The seed-based analysis was necessary because the winner-takes-all approach does not capture information about interactions among regions that fall in separate networks (Yeo et al., 2011). In contrast, the LDA estimate of cortical networks detects overlap among functional networks comprising early retinotopic visual areas and other components of this sensory-motor processing stream that were not previously appreciated (Fig. 4).

In particular, the estimated human LIP homologue is involved in all four networks in Fig. 4, including a network comprising mostly regions in the occipital lobe (network 1), a network comprising mostly regions in occipital and parietal lobes (network 2), a network comprising mostly regions in parietal and frontal lobes (network 3) and a higher association network spanning parietal, temporal and frontal lobes (network 4). Therefore networks 2 and 3 may serve as a bridge between early visual cortex (network 1) and higher association cortex (network 4). The FEF region is preferentially involved in the association networks (Networks 2–4), but less so in the early visual network (network 1), consistent with the fact that FEF occupies a higher position than LIP in the sensory-motor hierarchy (Fellman and Van Essen, 1991; Maunsell and Van Essen, 1983).

Parallel architecture, network interactions and human behavior

The neurophysiological literature has long recognized the importance of network interactions to cognition (e.g., Akam and Kullmann, 2010; Buschman et al., 2012; Buzsáki and Draguhn, 2004; Fries, 2005). In recent years there has been increasing interest to use human brain imaging to study how functionally distinct networks interact to support complex behaviors (Bassett et al., 2011; de Pasquale et al., 2012; Fornito et al., 2012; Spreng et al., 2010, 2013).

Our current work suggests that a particularly robust form of interaction might take place in association cortex (Figs. 2, 3, 4 and 8). This is consistent with anatomical evidence that heteromodal association regions serve as critical gateways for information processing (Goldman-Rakic, 1988; Jones and Powell, 1970; Mesulam, 1998; Pandya and Kuypers, 1969). While we have emphasized long-distance interactions by focusing on overlapping regions at some distance from network boundaries (Figs. 2b, 3 and 4), the interdigitated topography of networks in association cortex suggests a stereotypical spatial architecture, through which network components can locally interact with adjacent components of other functionally distinct networks (also see Power et al., 2013). In other words, local interactions among networks may occur within recurrent spatial patterns or motifs (Bullmore and Sporns, 2009; Power et al., 2011; Rubinov and Sporns, 2010; Sporns and Kotter, 2004). The expectation is that these local and distant interactions among functionally distinct networks facilitate emergent complex behaviors.

This network organization can be detected and replicated in individual subjects, although the details might vary across sessions and across subjects, especially in the association cortex (Figs. 5 and 6). This variability in network organization may underlie individual differences in behavior or even variability in behavior within a subject. This would be consistent with recent results showing that intersubject variability in functional connectivity is the highest in the association cortex, with regions predicting individual differences in cognitive domains predominantly located in regions of high connectivity variability (Mueller et al., 2013).

The distributed association networks converge on cortical regions that are late to develop (Hill et al., 2010) and are expanded in the human brain relative to the macaque brain (Van Essen and Dierker, 2007). Therefore, these distributed association networks have likely been under strong selective pressure to expand in recent hominin evolution. Consequently, even though this parallel architecture may also exist in monkeys (Goldman-Rakic, 1988), the details of the interdigitation and overlaps (Figs. 1, 3 and 4) may differ from monkeys and...
contribute to human cognition (see Buckner and Krienen, 2013 for discussion).

Limitations and future work

Tract tracing and physiological studies in monkeys and cats suggest that the cerebral cortex forms spatially organized circuits that include prominent connections with subcortical structures (for review, see Haber, 2003; Jones, 2007; Strick et al., 2009). In the present work we limit our analysis to the cerebral cortex, and therefore do not consider interactions between the cerebral cortex and subcortical structures. This limitation is partly methodological. Subcortical structures and the cerebellum can have significantly lower signal-to-noise ratio than the cerebral cortex and so are typically analyzed separately in functional connectivity studies (Di Martino et al., 2008; Zhang et al., 2008; Krienen and Buckner, 2009; Buckner et al., 2011; Choi et al., 2012; Dobromyslin et al., 2012; but see Habas et al., 2009; O’Reilly et al., 2010). Our (unpublished) observations suggest that including subcortical structures with cortical analyses would probably not alter the estimated cortical networks, even though subcortical results might be suboptimal. Consequently, we focus on the cerebral cortex in this paper. We are currently exploring techniques that can more effectively jointly analyze cortical and subcortical structures.

In addition, this work focuses on the spatial relationships among resting-state functional connectivity networks. The parallel architecture and overlaps in hierarchical and paralimbic networks suggest stereotypic spatial motifs in network interactions. However, our (essentially) correlational analyses will miss vital details of the dynamics of network interactions (Allen et al., in press; Chang and Glover, 2010; de Pasquale et al., 2012; Hutchinson et al., 2013; Smith et al., 2012). Our resting-state analyses will also miss changes in network interactions as a result of task performance (Bassett et al., 2006, 2011; Fornito et al., 2012; Spreng et al., 2010). These potential issues are outside the scope of this paper and will be addressed in future work.

The network organization we found is robust across datasets (GSP and HCP), preprocessing (with and without global signal regression), number of networks, degenerate/alternate network estimates and techniques (clustering, LDA and ICA). Although details vary, corroboration across the different analyses increases our confidence that the general organization of cortical networks is intrinsic to the data. However, we caution that the agreement among different techniques (clustering, LDA and ICA) may simply result from underlying assumptions common across the techniques. For example, while there are similarities between gradient-based (Cohen et al., 2008) and clustering/modularity-based parcellations (Power et al., 2011; Yeo et al., 2011), the boundaries in certain brain regions (e.g., visual cortex) can be qualitatively different (Wig et al., in press; Buckner and Yeo, submitted for publication). While the current approaches capture important features of cortical organization it is unclear to what degree they will overlap with gradient-based parcellation methods and what those differences might mean.

Conclusions

The human association cortex consists of multiple, interdigitated large-scale networks that, while partially overlapping, possess predominately parallel organization. This architecture can be detected and replicated in individual subjects. Many, but not all, association regions appear to participate in multiple networks, including those that lie some distance from estimates of network boundaries. The present work suggests that it is possible to consider both the divergent and convergent nature of connectivity in the human cerebral cortex using functional connectivity MRI. Our results suggest that segregation and interdigitation of networks in association cortex may be a true feature of cortical organization and not an artifact of the methods used to estimate their topography.

Acknowledgements

This work was supported by the National Medical Research Council, Singapore (SfA/0004/2008), Massachusetts General Hospital-University of California, Los Angeles Human Connectome Project (U54MH091165), the Simons Foundation and the Howard Hughes Medical Institute. Data were provided by the Brain Genomics Superstruct Project (Principal Investigators: Randy Buckner, Jordan Smoller, and Joshua Roffman) and by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; U54MH091657) 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. This research also utilized resources provided by the Center for Functional Neuroimaging Technologies, P41EB015869 and instruments supported by S10RR021110, S10RR023401 and S10RR019307 from the Athinoula A. Martinos Center for Biomedical Imaging at the Massachusetts General Hospital. The authors would like to thank Jessica Tandi for help with some of the figures in this paper.

Conflict of interest

There is no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2013.10.046.

References


Dosenbach, N.U.F., Fair, D.A., Miezin, F.M., Cohen, A.L., Wenger, K.K., Dosenbach, R.A.T.,

Buckner, R.L., Krienen, F.M., Yeo, B.T.T., 2013. Opportunities and limitations of functional

Fox, M.D., Zhang, D., Snyder, A.Z., Raichle, M.E., 2009. The global signal and observed

Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention

Corbetta, M., Patel, G., Shulman, G.L., 2008. The reorienting system of the human brain:

Mazoyer, N., Mazoyer, B., Mellet, E., et al., 2011. Brain activity at rest: a multiscale hi-


Mapp. 26, 15.


Lee, D.D., Seung, H.S., 1999. Learning the parts of objects by non-negative matrix factori-


