BRAIN CONNECTIVITY
Volume 9, Number 5, 2019

DOI: 10.1089/brain.2019.0668

A Novel Method for Extracting Hierarchical Functional Subnetworks Based on a Multisubject Spectral Clustering Approach

Xiaoyun Liang,[1,2] Chun-Hung Yeh,[1] Alan Connelly,[1,3] and Fernando Calamante[1,3,4]

Abstract

Brain network modularity analysis has attracted increasing interest due to its capability in measuring the level of integration and segregation across subnetworks. Most studies have focused on extracting modules at a single level, although brain network modules are known to be organized in a hierarchical manner. A few techniques have been developed to extract hierarchical modularity in human functional brain networks using resting-state functional magnetic resonance imaging (fMRI) data; however, the focus of those methods is binary networks produced by applying arbitrary thresholds of correlation coefficients to the connectivity matrices. In this study, we propose a new multisubject spectral clustering technique, called group-level network hierarchical clustering (GNetHiClus), to extract the hierarchical structure of the functional network based on full weighted connectivity information. The most reliable results of hierarchical clustering are then estimated using a bootstrap aggregation algorithm. Specifically, we employ a voting-based ensemble method, that is, majority voting; random subsamples with replacement are created for clustering brain regions, which are further aggregated to select the most reliable clustering results. The proposed method is evaluated over a range of group sample sizes, based on resting-state fMRI data from the Human Connectome Project. Our results show that GNetHiClus can extract relatively consistent hierarchical network structures across a range of sample sizes investigated. In addition, the results demonstrate that GNetHiClus can hierarchically cluster brain functional networks into specialized subnetworks from upper-to-lower level, including the high-level cognitive and the low-level perceptual networks. Conversely, from lower-to-upper level, information processed by specialized lower level subnetworks is integrated into upper level for achieving optimal efficiency for brain functional communications. Importantly, these findings are consistent with the concept of network segregation and integration, suggesting that the proposed technique can be helpful to promote the understanding of brain network from a hierarchical point of view.

Keywords: bootstrapping; functional connectivity; hierarchical clustering; spectral clustering

Introduction

Modularity of brain networks provides selective adaptability, which has been demonstrated to play a key role in developmental optimization of the human brain (Meunier et al., 2009a). Hence, the detection of such community structures within brain networks, namely communities or modules, has attracted increasing interest (Ferrarini et al., 2009; Meunier et al., 2009b). Methods have been proposed to seek the optimal modules for a variety of complex networks (i.e., not limited to brain networks); a fundamental characteristic of such ''modular'' system is that nodes within modules are densely connected, whereas cross-connections between modules are sparser (Newman and Girvan, 2004; Newman, 2006).

Importantly, Simon proposed that most complex systems are organized in a *hierarchical* manner, introducing the notion of ''nearly-decomposable systems'' (Simon, 1962), which has been well accepted as fundamental for cognitive and computational neuroscience. In a stricter sense, a system is decomposable if its function can be decomposed to the sum of the independent functions of its parts. Furthermore,

[1]The Florey Institute of Neuroscience and Mental Health, Heidelberg, Australia.
[2]Mary Mackillop Institute for Health Research, Australian Catholic University, Melbourne, Australia.
[3]The Florey Department of Neuroscience and Mental Health Medicine, University of Melbourne, Melbourne, Australia.
[4]Sydney Imaging, School of Aeronautical, Mechanical and Mechatronic Engineering, The University of Sydney, Sydney, Australia.

Simon suggested that hierarchies reduce the complexity of systems, which supports the feature of modularity. In this regard, a module (or a subsystem) can be considered as a self-enclosed unit that works almost independently, regardless of the changes that occur in other modules. Human brain network is an instantiation of such a hierarchically decomposable system.

Hierarchical modular networks

Previous studies support the concept that a brain functional network has an intrinsic hierarchical structure (Moretti and Munoz, 2013). However, most previous studies focused on studying modularity at a single level of the community structure, without mapping the properties or functions of the submodular communities at other hierarchical levels (Chen et al., 2008; Meunier et al., 2009a). Extracting hierarchical structures from the brain network should help gain deeper insights and advance our understanding of brain organization from a network point of view.

Weighted versus binary networks

There have been few studies developed to extract the hierarchical subnetworks. A recent study investigated the human functional brain networks at several hierarchical levels by applying a method for extracting hierarchical subnetworks (Meunier et al., 2009b). In another study, an approach was proposed to extract hierarchical functional modularity using an unbiased clustering coefficient (Ferrarini et al., 2009). Hierarchical functional modularity was also investigated in two more recent studies (Bassett et al., 2010; Power et al., 2011). Notably, all these studies require setting an arbitrary threshold for the correlation coefficient for measuring functional connectivity strength; different thresholds would unfavorably lead to distinct clustering outcomes. Furthermore, although binary networks obtained by applying arbitrary thresholds remain the most common approach employed for network analysis, weighted networks are increasingly recognized as they carry continuous characterization of brain connectivity (Bassett and Bullmore, 2017). Applying arbitrary thresholds could impose a detrimental effect on the subsequent network analyses (Garrison et al., 2015); a binarized connectivity matrix that discards putative connectivity strength might further deviate network measures from realistic biological connectivity, leading to compromised network analysis results.

Group-level versus individual-level networks

In general, functional networks are analyzed at either individual level or group level. Although the former could be important for some applications, such as subject-specific treatment (so-called precision medicine), its use has been limited due mainly to the relatively low signal-to-noise ratio of functional magnetic resonance imaging (fMRI) data. Therefore, group-level network analysis has been most commonly employed. Brain functional networks at group level are of great interest regarding probing the mechanism of complex brain function, and how networks change across disease condition and mental states. Group-level analyses have been mostly performed through the following ways: (1) direct averaging of individual-level networks across the group to obtain a single representative network for the group (Achard et al., 2006; Liang et al., 2014) or (2) estimating group-level networks by accounting for intersubject variability with machine learning techniques (Liang et al., 2016, 2018; Ng et al., 2013; Varoquaux et al., 2010). Despite the popularity of the former approach, it fails to consider the variation among individual subjects. Thus, the latter studies advocate the incorporation of individual-level constraint to increase their robustness.

In this study, our objective is to robustly identify functional subnetworks of multiple hierarchical levels, directly from full weighted networks without filtering networks using arbitrary thresholds. To achieve this aim, we propose a novel multisubject spectral clustering method in conjunction with a bootstrap ensemble learning technique, called group-level network hierarchical clustering (GNetHiClus). Furthermore, the performance of GNetHiClus is assessed in terms of robustness and reproducibility based on resting-state fMRI data downloaded from the Human Connectome Project (HCP).

Materials and Methods

Motivation for developing GNetHiClus

There have been a variety of clustering algorithms available, of which the spectral clustering approach is very useful in hard nonconvex clustering problems (Donath and Hoffman, 1973). Spectral clustering views the data clustering as a graph partitioning problem without relying on any assumption on the form of data clustering. As such, the clustering results obtained by spectral clustering generally are more accurate than other approaches. Furthermore, spectral clustering can be easily implemented by applying linear algebra methods. Overall, therefore, spectral clustering provides a very efficient approach for clustering.

For conventional spectral clustering, *k*-means clustering is further applied to the eigenvectors obtained with spectral clustering, to determine the correct number of clusters. However, *k*-means clustering could be sensitive to initial conditions; subsequent analyses might be unfavorably affected by such uncertainty. Furthermore, such standard spectral clustering approaches are limited to extracting modularity at one particular level while ignoring hierarchical network topology.

To address the abovementioned limitations, we propose a hierarchical spectral clustering approach based on recursive biclustering. Specifically, subnetworks at the next level are obtained by subdividing each subnetwork at the current level, thereby providing detailed hierarchical network topology across multiple levels. Importantly, each biclustering process is exclusively based on spectral clustering, without being subject to the uncertainty issues intrinsic to k-means clustering.

In this study, we extend the multiview spectral clustering technique (Kumar et al., 2011) to a multisubject scenario, which could effectively integrate useful intersubject information from every individual subject within a group through regularization, thereby achieving more reliable clustering results. By combining the hierarchical biclustering and multiview spectral clustering techniques, we propose the *GNetHiClus* method to identify functional subnetworks at hierarchical level by sharing individual-network information across subjects.

The spectral clustering technique

In general, spectral clustering is implemented as follows:

1. Adjacency matrix: For a given set of data points $x_1, \ldots, x_n$, a similarity graph $G=(V_t, E)$ can be employed to represent the data, with $V_t = \{v_1, \ldots, v_n\}$ and E representing vertices and edges, respectively. The adjacency matrix $A=\left(a_{i,j}\right)_{i,j=1,\ldots,n}$ of the graph is employed to characterize the similarity between vertices v_i and v_j, with a non − negative *weight* a_{ij} .
2. Normalized adjacency matrix (Chung, 1997)

$$N_1 = D^{-1/2} A D^{-1/2}, \tag{1}$$

where the degree matrix D is defined as diagonal matrix with node degrees $d_1, \ldots, d_n$ on the diagonal. To make sure normalized adjacency matrix is symmetric, we let $N=(N_1+N_1{}^T)/2$. Spectral clustering algorithm solves the following optimization problem for the normalized adjacency matrix N:

$$\underset{U \in R^{n\times k}}{max}\ tr\left(U^T N U\right), \quad \text{s.t. } U^T U = I,$$

where tr represents the trace of a matrix.

3. Compute k eigenvectors corresponding to the k largest eigenvalues. Specifically, here we consider $k=2$ (with two eigenvectors u_1 and u_2) because our proposed approach relies on iterative biclustering. In particular, the second largest eigenvector u_2 is also called a Fiedler vector, which can be used to divide its elements into two clusters based on their associated signs (Fiedler, 1973), that is, a cluster corresponding to the data points with elements of "+" sign and another cluster with those that have "−" sign (see Fig. 1 for an illustrative example).

The original multiview spectral clustering approach

Basically, a multiview spectral clustering approach can be considered as a way to extract information from different data representations (such as multimodality or multisubject data) to achieve more accurate clustering results. Although multiple views of data might be available in many scenarios, such complementary information cannot be fully utilized without an appropriate multiview approach. To fully utilize such complementary information (i.e., information across subjects in our study), a coregularized multiview spectral clustering was proposed by Kumar et al. (2011); they demonstrated the advantages of their approach over other existing techniques for data clustering. In that study, a centroid-based coregularization was proposed for regularizing each view-specific set of eigenvectors $U^{(v)}$ $(v=1,2,\ldots,m)$ toward a common centroid $U^{(*)}$, with m representing the total number of views. The objective function is formulated as

$$\underset{U^{(1)}, U^{(2)}, \ldots, U^{(m)}, U^* \in R^{n\times k}}{\max} \sum_{v=1}^{m} [tr\left(U^{(v)^T} N^{(v)} U^{(v)}\right) + \lambda_v tr\left(U^{(v)} U^{(v)^T} U^* U^{*^T}\right)], \tag{2}$$

$$\text{s.t. } U^{(v)^T} U^{(v)} = I,\ \forall 1 \le v \le m,\ U^{*^T} U^* = I.$$

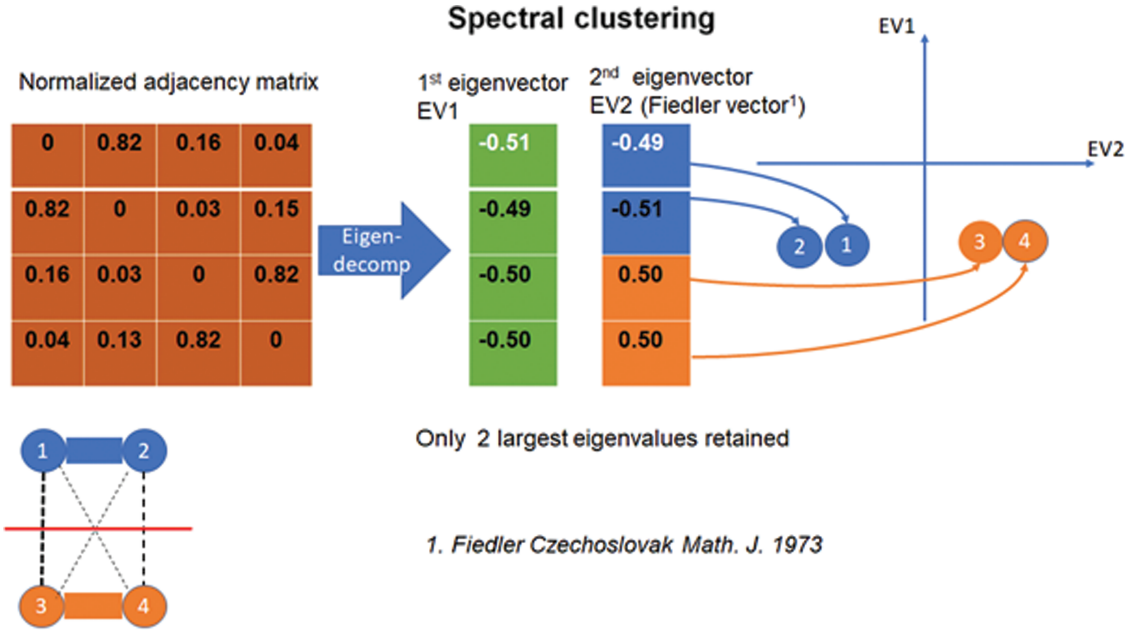

FIG. 1. An illustrative example of the overall concept of spectral clustering. Given a normalized adjacency matrix with four nodes, Fiedler vector (i.e., the second eigenvector derived from eigen decomposition of the input matrix) can be employed to divide the original four nodes (1, 2, 3, and 4) into two clusters based on the positive and negative signs of the values of Fiedler vector. In this example, nodes 1 and 2 are classified into a cluster, and nodes 3 and 4 are assigned into another. Color images are available online.

This objective function essentially tries to estimate consensus eigenvectors $U^{(*)}$ from m views through regularization; each regularization term is weighted by a weighting parameter λ_v that reflects the importance of view v.

The objective function can be solved in an alternative way that optimizes one single view-specific eigenvector $U^{(v)}$ at a time, while fixing all other variables. Thus, the objective function for optimizing $U^{(v)}$ for view v is reformulated as follows:

$$\max_{U^{(v)}\in R^{n\times k}} tr\left(U^{(v)^T}N^{(v)}U^{(v)}\right)+\lambda_v tr\left(U^{(v)}U^{(v)^T}U^*U^{*^T}\right), \quad \text{s.t. } U^{(v)^T}U^{(v)}=I \tag{3}$$

Equation (3) can be written as follows:

$$\max_{U^{(v)}\in R^{n\times k}} tr\left(U^{(v)^T}(N^{(v)}+\lambda_v U^*U^{*^T})U^{(v)}\right), \text{s.t. } U^{(v)^T}U^{(v)}=I. \tag{4}$$

This provides a standard spectral clustering objective for $U^{(v)}$ with a modified normalized adjacency matrix $N^{(v)}+\lambda_v U^*U^{*^T}$.

Next, consensus $U^{(*)}$ is optimized by keeping all view-specific eigenvectors $U^{(v)}$ fixed. With Equation (2), $U^{(*)}$ can be obtained by solving the following equation:

$$\max_{U^*\in R^{n\times k}} \sum_v \lambda_v tr\left(U^{(v)}U^{(v)^T}U^*U^{*^T}\right), \text{s.t. } U^{(*)^T}U^{(*)}=I. \tag{5}$$

The cyclic property of matrix traces (i.e., for square matrices A, B, and C, $Tr(ABC)=Tr(BCA)=Tr(CAB)$) is then applied to reformulate Equation (5) as a standard spectral clustering objective for $U^{(*)}$ as follows:

$$\max_{U^*\in R^{n\times k}} tr\left(U^{*^{(T)}}\left(\sum_v \lambda_v\left(U^{(v)}U^{(v)^T}\right)\right)U^*\right), \text{s.t. } U^{(*)^T}U^{(*)}=I, \tag{6}$$

with a modified normalized adjacency matrix $N_m=\sum_v \lambda_v\left(U^{(v)}U^{(v)^T}\right)$.

Equation (6) could be reformulated as $\max_{U^*\in R^{n\times k}} tr\left(U^{*^{(T)}}U^*\right)$, s.t. $U^{(*)^T}U^{(*)}=I$; the solution to this equation can be obtained by using the standard spectral clustering approach (Ng et al., 2002), given the provided standard spectral clustering objective for $U^{(v)}$ $(v=1,2,\ldots,m)$ and $U^{(*)}$. In practice, the multiview spectral clustering approach is iteratively implemented; the clustering procedure stops when the difference in value of objective function between consecutive iterations falls below a relatively low threshold $(\varepsilon=10^{-4})$ (Kumar et al., 2011), suggesting that the optimal solution is achieved.

The proposed GNetHiClus approach

GNetHiClus consists of two parts (parts I and II)—A flowchart is shown in Figure 2:

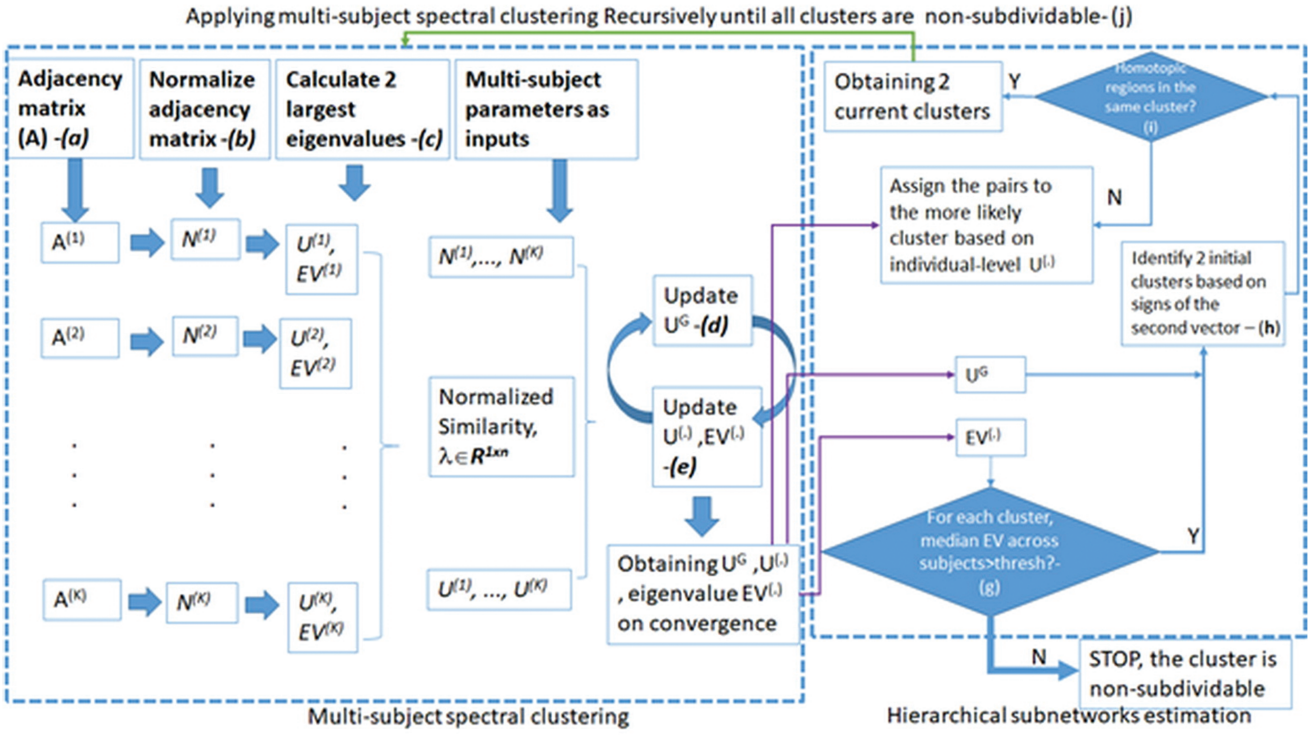

FIG. 2. The workflow of GNetHiClus for extracting hierarchical functional subnetworks. The flowchart can be divided into two major parts, as grouped by two dashed boxes: Left: multisubject spectral clustering [steps (a) to (f)]; Right: Hierarchical clustering [steps (g) to (j)]. See Materials and Methods section for the detailed description of each step and the definition of each symbol within the flowchart. Note: this procedure needs to be repeated for a certain number of times (e.g., 100 times in this study) to realize ensemble learning through bootstrapping. GNetHiClus, group-level network hierarchical clustering. Color images are available online.

(I) *The multisubject spectral clustering already* described is used to estimate a consensus (group level) eigenvector of normalized adjacency matrices and consistent individual-level eigenvectors and eigenvalues across the group. It includes the following steps:

(a) Using blood oxygenation level-dependent (BOLD) fMRI data to generate an adjacency matrix $A^{(v)}$ for a subject v ($v = 1,\ldots,m$), where m is the number of subjects within the group.
(b) Calculating normalized adjacency matrix $N_1{}^{(v)} = D^{(v)\,-1/2} A^{(v)} D^{(v)\,-1/2}$, D: degree matrix $N = (N_1 + N_1{}^T)/2$.
(c) Calculating eigenvectors $U^{(v)}$ by eigenvalue decomposition of $N^{(v)}$. In particular, we are interested in the second eigenvalue and eigenvector (see (h) hereunder).
(d) Updating group consensus U^G: solving $\max_{U^G \in R^{n\times k}} trU^{G^T}(\sum_v \lambda_v \left(U^{(v)} U^{(v)^T}\right))U^G\}$, s.t. $U^{G^T} U^G = I$, by eigenvalue decomposition on $\sum_v \lambda_v \left(U^{(v)} U^{(v)^T}\right)$, with $U^{(v)}$ from (c) for first iteration, *n:* number of nodes, *k:* number of clusters, λ_v: weight (i.e., the ''importance'' of subject v, $\lambda_v = 1 - D$, with D representing disagreement between subjects (views) (see Kumar et al., 2011, for the calculation of D (Eq. 2); note: $k = 2$ due to biclustering in our study, therefore, $k = 2$ for GNetHiClus).
(e) Updating $U^{(v)}$: solving $\max_{U^{(v)} \in R^{n\times k}} tr\left\{U^{(v)^T}(N^{(v)} + \lambda_v U^G U^{G^T})U^{(v)}\right\}$, s.t. $U^{(v)^T} U^{(v)} = I$, by eigenvalue decomposition on $N^{(v)} + \lambda_v U^G U^{G^T}$, with all other $U^{(v)}$ and U^G being fixed.
(f) Updating $U^{(v)}$ and U^G iteratively through (d) and (e) until convergence (Kumar et al., 2011), yielding final estimates of U^G, $U^{(v)}$, and $EV^{(v)}$ (i.e., eigenvalues).

(II) *The hierarchical clustering* is used to hierarchically divide the networks (or subnetworks) into two clusters downstream from the current level using U^G, $U^{(v)}$, and $EV^{(v)}$ from (I).

(g) For each cluster, if the median of Fiedler eigenvalues of all individual subjects across the group is positive (i.e., more than half of the subjects have a positive second largest eigenvalue), continue to step (h); otherwise, stop performing further clustering.
(h) The nodes are assigned into two different clusters based on the signs of the Fiedler vectors (i.e., ''+'' or ''−'').
(i) Furthermore, since brain subnetworks tend to be bilateral (Meunier et al., 2009b; Tyszka et al., 2011), homotopic brain regions are forced to be present in the same subnetworks to which the homotopic regions are more likely belonging, which is determined as follows:

1. For region i (and j), whether i (and j) is identified in cluster A or cluster B is first determined at individual level, that is, total number of subjects that vote for region i (and j) to be in clusters A and B are counted.
2. Then, for region i (and j), the magnitudes of eigenvector entries corresponding to region i are summed across those subjects identifying i (and j) to be in cluster A at individual level, that is, $S_A(i)$ and $S_A(j)$ are computed.
3. Similarly, for region i (and j), the magnitudes of eigenvector entries corresponding to region i are summed across those subjects identifying i (and j) to be in cluster B at individual level, that is, $S_B(i)$ and $S_B(j)$ are computed.
4. Finally, if $S_A(i) + S_A(j) > S_B(i) + S_B(j)$, then the (i,j) pair is assigned to cluster A, otherwise, it is assigned to cluster B.

(j) Two new clusters obtained at the current level are fed into (I) recursively, until all clusters cannot be divided based on the criteria in (g).

Matlab code to perform the proposed method can be downloaded from (https://www.florey.edu.au/science-research/scientific-services-facilities/human-mri/imaging-software).

MRI data acquisition

Resting-state FMRIB's ICA-based X-noiseifier-denoised (Salimi-Khorshidi et al., 2014) fMRI data of 133 subjects ($n = 133$) were downloaded from the ConnectomeDB. Data of 1200 volumes had been acquired in a Siemens 3T Connectome Skyra system, using a gradient-echo slice-accelerated multiband echo planar imaging sequence at 2 mm isotropic resolution, with TR/TE = 720/33.1 msec. See Glasser et al. (2016) and Van Essen et al. (2016) for HCP data acquisition details.

Data analysis

Network construction. Networks are constructed as follows: (1) the automated anatomical labeling (AAL) template (Tzourio-Mazoyer et al., 2002) was used to parcellate the whole brain into 90 regions (cerebellum was excluded); (2) region-wise mean time series were calculated across all voxels within each region (size: 90×1200); (3) individual-level functional connectivity matrix (size: 90×90) was constructed from region-wise mean time series of each subject using Pearson correlation.

Hierarchical clustering results: statistical modeling and analyses. GNetHiClus is implemented with the following two-level analysis. First, the first-level analysis for selecting most representative clustering results across bootstrap data sets. Second, the second-level analysis for identifying the final clustering results with the highest probability across repetitions of the first-level analysis.

(1) Bootstrapping: Since *GNetHiClus* is implemented across all subjects, only a single group outcome is obtained, to which regular statistical analyses could not be applied. To address this issue, the bootstrap method relying on random subsampling with replacement was employed. For a given number of total subjects (i.e., a subset of the maximum 133 subjects, which was considered to investigate the effect of sample size), random subsampling was implemented using a fixed ratio of $\sim 60\%$ with a variable subsampling size (i.e., the ratio between the number of randomly subsampled

subjects and the total number of subjects in the pool): 75/125, 70/117,65/108,60/100,55/92, and 50/83. As an example, for the case of 75/125, each subsampled data set was obtained by randomly selecting 75 subjects out of a total of 125 subjects; this procedure was then repeated 100 times to generate 100 subsampled data sets. We sought to identify the optimal implementation strategies for the bootstrapping among all these six cases, that is, we aim to determine the optimal subsamples (and total subjects) that can achieve robust hierarchical clustering results. As more reliable results are expected when using larger sample sizes, the clustering results of 80/133 are employed as a ''ground-truth'' for evaluating the other results with smaller sample sizes.

(2) Extraction of the most representative clustering results for each of the six cases: (a) GNetHiClus was applied to each of the 100 bootstrap subsamples to extract hierarchical subnetworks; (b) normalized mutual information (NMI) (Fred and Jain, 2003) was calculated between each subsample and the remaining 99 subsamples, yielding a 100×99 matrix of NMI values; (c) the 99 NMI values were averaged, yielding a vector of 100×1 mean NMI (mNMI) values; (d) the subsample having the highest mNMI was selected from the vector of 100 mNMIs; (e) hierarchical subnetworks that produced the highest mNMI were designated as the most reliable clustering results. The reasons for choosing NMI to evaluate the proposed method are as follows: (a) it is a good measure for determining the performance of clustering and (b) NMI can be used to compare clustering results having different number of clusters (Danon et al., 2005). In general, the higher NMI values indicate that the more accurate clustering results are obtained (NMI$\in [0, 1]$), with NMI$= 1$ indicating perfect clustering results.

(3) Measurement of probability of clusters: (a) step (2) was repeated 100 times to obtain 100 groups of clusters (i.e., 10,000 bootstraps), of which, each group was expressed as a 90×90 matrix and all members of each cluster were denoted with a unique number; (b) for each of the 100 groups (i.e., 100 matrices of size 90×90), we counted the total number of occurrences of each cluster, $T_{cluster}$, by searching over all clusters and all groups; (c) for each of the 100 groups, the total number of occurrences of all clusters, T_{all}, belonging to the group was calculated; (d) the group with the maximum T_{all} was then selected as the optimal group, which yielded the clustering results with the highest overall probability; (e) those clusters corresponding to the identified optimal group were designated as the final clustering results; (f) for the obtained clustering results, the probability of each cluster, p, was calculated as $p = T_{cluster}/100$.

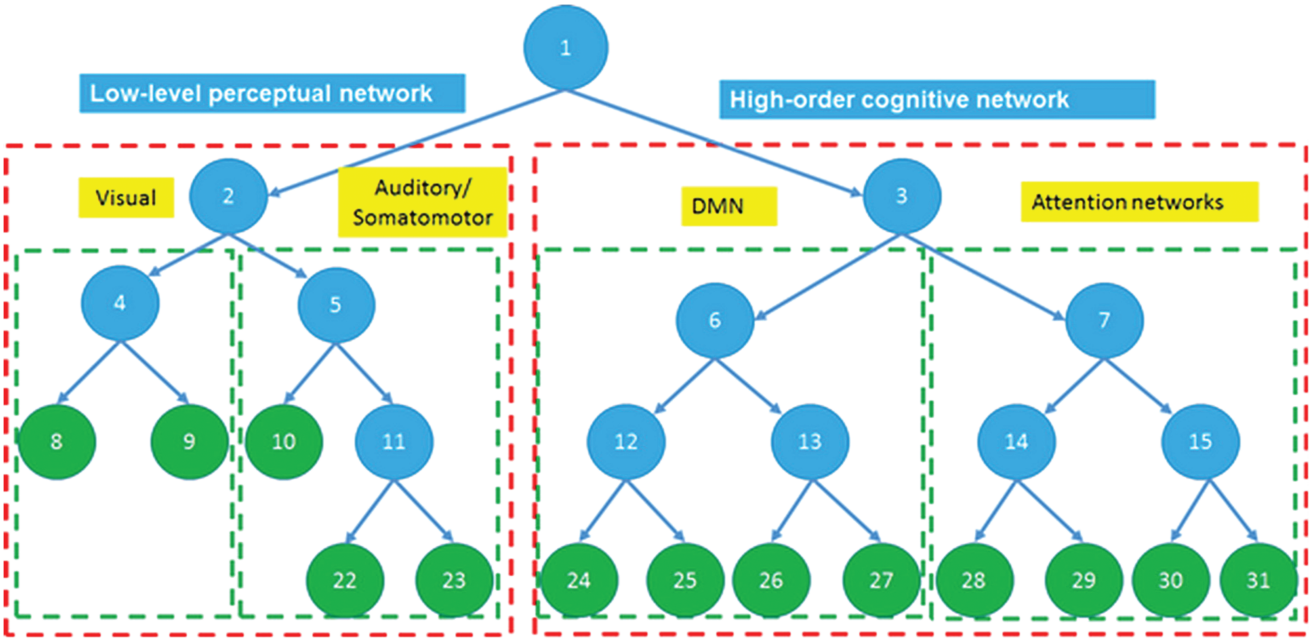

FIG. 3. Results of hierarchically clustered subnetworks using GNetHiClus, for the case of using 75/125 subsample/sample size. From upper-to-lower level of the figure, the brain network is hierarchically subdivided into two clusters until stopping criterion is met. Each circle indicates a cluster, with green clusters representing the final clusters and blue clusters representing clusters at intermediate stage. At the second level (i.e., following the first division), the result shows that GNetHiClus subdivides the whole brain functional network into two meaningful major subnetworks, one corresponds to the low-level perceptual network (cluster 2) and the other one corresponds to the higher order cognitive network (cluster 3). At the third level, the result shows that the visual network and the SMN are separated from low-level perceptual network, whereas the DMN and the attention network are divided by the high-order cognitive network. These functional specialized subnetworks can be further divided into subnetworks at finer levels (fourth and fifth) to fulfill specialized tasks. Likewise, from lower-to-upper level, the outcome of hierarchical structure can be used to conversely integrate subnetworks into whole-brain network. Also refer to Figure 4 and Table 1 for a detailed description of the constituents of each subnetwork. DMN, default-mode network; SMN, somatomotor network. Color images are available online.

Results

Our results show that GNetHiClus extracts 13 functional subnetworks (the green clusters at the bottom level) from resting-state HCP fMRI data for subsample/sample sizes of 75/125 (Fig. 3). Importantly, the results suggest that GNetHiClus can automatically organize the whole brain functional networks into a hierarchical ''tree'' structure, where each branch represents a well-established functional subnetwork of the human brain. As shown in Figure 3, the whole-brain network is first divided into *low-level perceptual* and *high-order cognitive* networks, from which well-known resting-state brain networks are hierarchically clustered; these subnetworks include visual network, auditory/somatomotor network (SMN), default-mode network (DMN), and attention network.

Specifically, 13 functional subnetworks extracted at the finest levels, that is, fourth and fifth levels shown in Figure 3, are shown in Figure 4. The nodes of these undividable subnetworks are displayed using the AAL atlas. To facilitate interpretation, the undividable subnetworks are aggregated (but colored differently to differentiate the clusters) into four well-known resting-state networks, that is, visual network, auditory/SMN, DMN, and attention network. Table 1 summarizes the brain nodes of each cluster for the case of using subsamples/samples = 75/125,70/117, 60/100, and 55/92 showing high correspondence between those clustered brain regions with their roles known in human brain functions.

As expected, increasing subsample/sample sizes improved clustering results. For 70/117 and 75/125, the clustering results were consistent with 80/133, as shown in Figure 5a, also achieving more persistent probabilities of all clusters (Fig. 6e, f) than all other cases (Fig. 6a–d). For lower subsampling/sampling sizes, that is, 65/108, slightly different clustering results were obtained: the bilateral dorsal cingulate gyrus, which were classified into auditory network/SMN for 80/133, were erroneously identified as nodes of attention networks

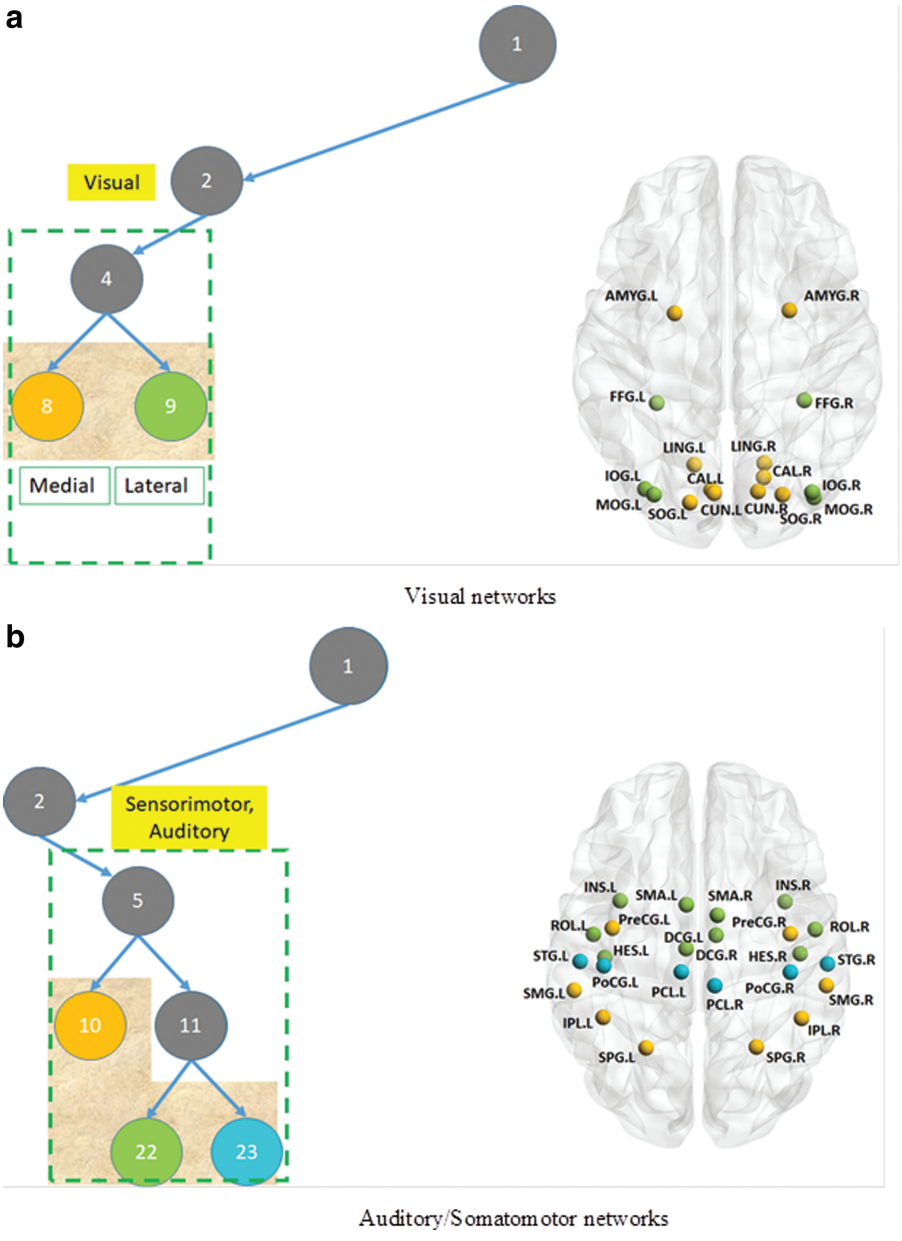

FIG. 4. Visualization of all functional subnetworks at the finest levels, that is, fourth and fifth levels shown in Figure 3. The whole-brain network is divided into four meaningful resting-state functional networks, that is, visual network, auditory/SMN, DMN, and attention network, which are shown in Figure 3 **(a–d)**, respectively. In each subfigure, those colored circles compose the current focused network. The four resting-state networks can be further partitioned into finer subnetworks (until indivisible), with each subnetwork represented in a different color. See Table 1 for definition of node acronyms. We just show the relevant part of the tree in each subfigure (see Fig. 3 for the full tree), with each colored circle corresponding to a set of nodes with the same color mapped onto the brain; as an example, **(a)** focuses on visual network, where yellow circle represents medial visual network and green circle represents lateral visual network. Color images are available online. *Continued.*

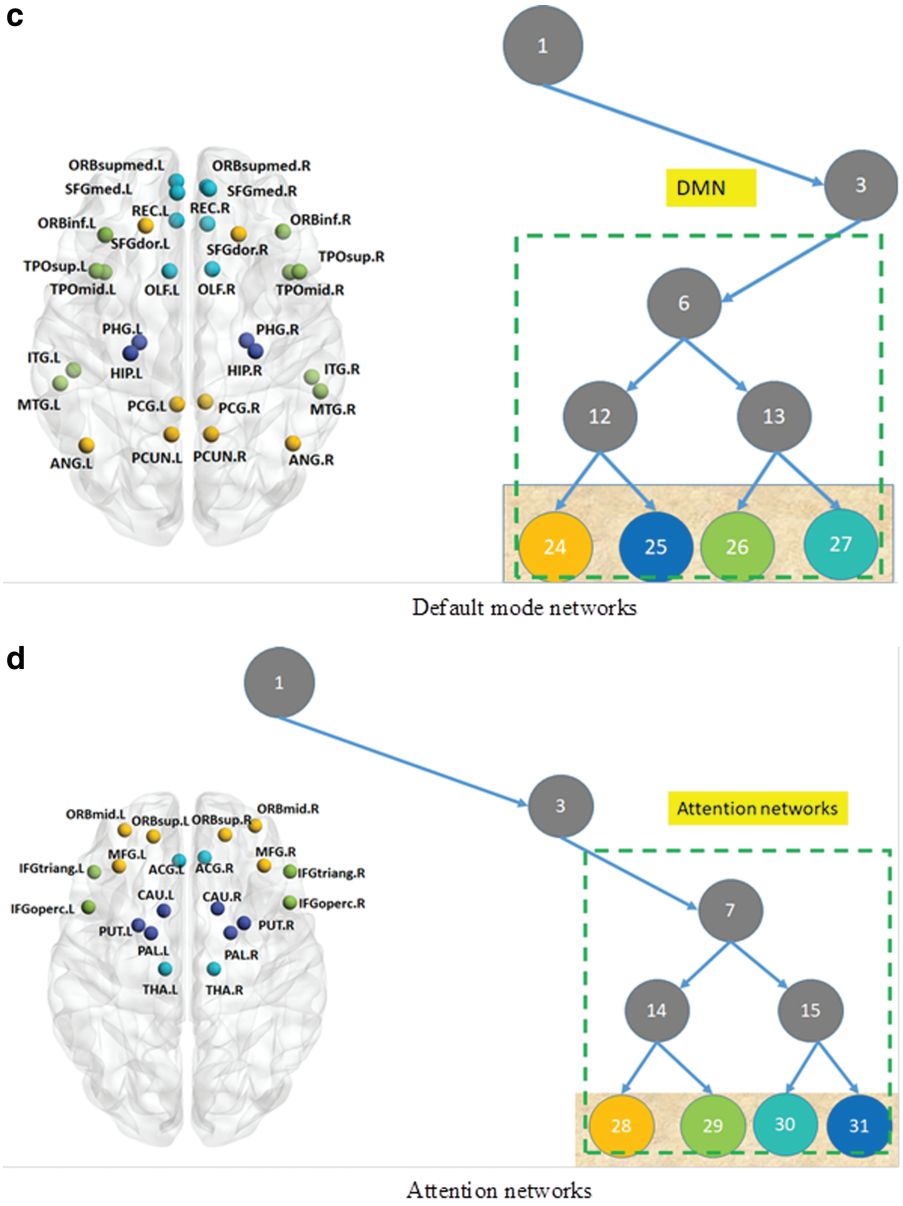

FIG. 4. (Continued).

(Fig. 5b). Although clustering results of 60/100 and 55/92 were also consistent with those of 80/133, these probabilities are not persistent across all clusters, that is, some clusters have very low probabilities (Fig. 6b, c). As expected, when using the lowest subsampling/sampling size of 50/83, in addition to the common difference observed in 65/108, one of the subnetworks (i.e., cluster 26 in Fig. 3) within DMN was further divided into two clusters, showing apparent deviation from the ground truth of 80/133 (see Fig. 5c; cluster 1: ORBinf.L, ORBinf.R, ITG.L, and ITG.R; cluster 2: MTG.L, MTG.R, TPOsup.L, TPOsup.R, TPOmid.L, and TPOsup.R).

With the clustering results from 80/133 as ground truth, overall high NMI values are obtained. For clusters having the highest overall probability among all 100 repetitions of each case, their NMI values were 0.9662, 1, 1, 0.9806, and 1, 1 for sampling sizes of 50/83, 55/92, 60/100, 65/108, 70/117, and 75/125, respectively. This demonstrates that GNetHiClus achieves clustering outcomes with fairly high accuracy (as indicated by high NMI values) for most of the clusters even with relatively low subsampling/sampling sizes, that is, 50/83.

Discussion

In this study, a multisubject hierarchical spectral clustering method, GNetHiClus, was proposed. Specifically, this is a recursive biclustering approach, where brain nodes are hierarchically biclustered into subnetworks until stopping criteria are met. In this way, functional hierarchical network structure of the human brain can be reconstructed from resting-state fMRI data. A bootstrap ensemble learning technique is adopted to compute the most representative functional hierarchical structure by repeatedly applying GNetHiClus on the bootstrap data sets (i.e., a subset of randomly selected

Table 1. Brain Regions of Subnetworks That Are Consistently Extracted from Human Connectome Project Data for Both Cases: 75/125, 70/117, 60/100, and 55/92 Using Group-Level Network Hierarchical Clustering (Fig. 5A)

Cluster No.	*Brain regions in each subnetwork*
Cluster 8	Amygdala (AMYG), calcarine (CAL), cuneus (CUN), lingual gyrus (LG), superior occipital gyrus (SOG)
Cluster 9	Middle occipital gyrus, inferior occipital gyrus, fusiform gyrus
Cluster 10	Precentral gyrus (PreCG), superior parietal gyrus (SPG), inferior parietal gyrus (IPG), supramarginal gyrus (SMG)
Cluster 22	Rolandic operculum (ROL), supplementary motor area (SMA), insula (INS), middle cingulum (DCG), heschl gyrus (HES)
Cluster 23	Postcentral gyrus (PoCG), paracentral lobule (PCL), superior temporal gyrus (STG)
Cluster 24	Superior frontal gyrus (SFG), posterior cingulum (PCG), angular gyrus (ANG), precuneus (PCUN)
Cluster 25	Hippocampus (HIP), parahippocampal gyrus (PHG)
Cluster 26	Inferior frontal–orbital part (ORBinf), temporal pole–superior part (TPOsup), middle temporal gyrus (MTG), temporal pole–middle part (TPOmid), inferior temporal gyrus (ITG)
Cluster 27	Olfactory (OLF), superior frontal–medial part (SFGmed), medial frontal–orbital part (ORBsupmed), rectus (REC)
Cluster 28	Superior frontal–orbital part (ORBsup), middle frontal gyrus (MFG), middle frontal gyrus–orbital part (ORBmid)
Cluster 29	Inferior frontal gyrus–opercular part (IFGoperc), inferior frontal gyrus–triangular part (IFGtriang)
Cluster 30	Anterior cingulum (ACG), thalamus (THA)
Cluster 31	Caudate (CAU), putamen (PUT), pallidum (PAL)

Cluster No. corresponds to the number of indivisible clusters labeled with green (i.e., the final clustering outcomes) shown in Figure 3. All brain regions are bilaterally present in the related clusters.

individuals). The performance of the proposed method in clustering brain networks is evaluated by using the publicly available HCP data; the consistency of our results with the concept of low-level perceptual and high-order cognitive networks (Mottron et al., 2006; Shang et al., 2014) along with hierarchical subnetworks (such as visual network, SMN, DMN, and attention network) demonstrates that the proposed method can achieve biologically reasonable hierarchical structures that characterize specialized functional networks. The main novelties of GNetHiClus are discussed point-by-point in the following sections:

(1) Avoiding uncertainty of choosing arbitrary thresholds: Previous network clustering approaches commonly rely on applying arbitrary thresholds to the correlation coefficients of inter-regional functional connectivity, with different thresholds possibly leading to different network clustering outcomes (or modularity). Empirically, the computation of network modularity is highly dependent on network density, with higher modularity achieved by using lower network density; as an extreme case, network modularity analysis might not be practically feasible if no thresholds are applied (i.e., for fully connected networks, conventional network modularity might detect the whole brain network as a single modular). In contrast, GNetHiClus takes weighted networks as the input without requiring arbitrary thresholds and binarizing the networks; it is rather immune to network density because it directly uses full networks for performing clustering, where per edge connectivity strength is the main determinant to the outcome of network clustering. Therefore, GNetHiClus can exclude the uncertainty induced by the selection of threshold values, potentially providing a more robust and consistent approach for investigating brain networks. However, direct comparisons between GNetHiClus and the existing threshold-based techniques are beyond the scope of this study.

(2) Ensemble learning to reduce sensitivity to noise: Network clustering is prone to the influence of noise, particularly in resting-state functional connectivity where BOLD contrast is sensitive to noise; it is, therefore, not ideal to rely on a single clustering result. Aggregation of many estimators could reduce the effect of noise, leading to more accurate results. In this study, bootstrapping is employed to generate randomly subsampled data sets with replacement: we employed an ensemble learning approach that is analogous to a majority voting-based ensemble algorithm (Lam and Suen, 1997); it trains models through randomly selecting multiple sets of subsamples from the training data sets, and then aggregates the votes (i.e., clustering results) from different models to decide the final class of the test object. Specifically, GNetHiClus is implemented with two-level analysis, that is, the first-level analysis for selecting most representative clustering results across 100 bootstrapping data sets and the second-level analysis for identifying clustering results with highest probability across first-level 100 repetitions. In contrast, previous methods have been mainly relying on obtaining average connectivity matrix from a group followed by the identification of (hierarchical) subnetworks. In particular, we proposed a novel and robust machine learning framework for investigating hierarchical networks, which can quantify the probability of extracting hierarchical functional modularity.

(3) Uncovering hierarchical structure: Although there is no direct ground truth available, the identified subnetworks of our hierarchical structure are overall consistent with resting-state networks extracted using other approaches, such as independent component analysis (ICA) (Beckmann et al., 2005); note, however, that ICA cannot yield hierarchical network structure. Previous studies have shown that visual network, sensorimotor network, auditory network, DMN, attention network, executive control network, and subcortical

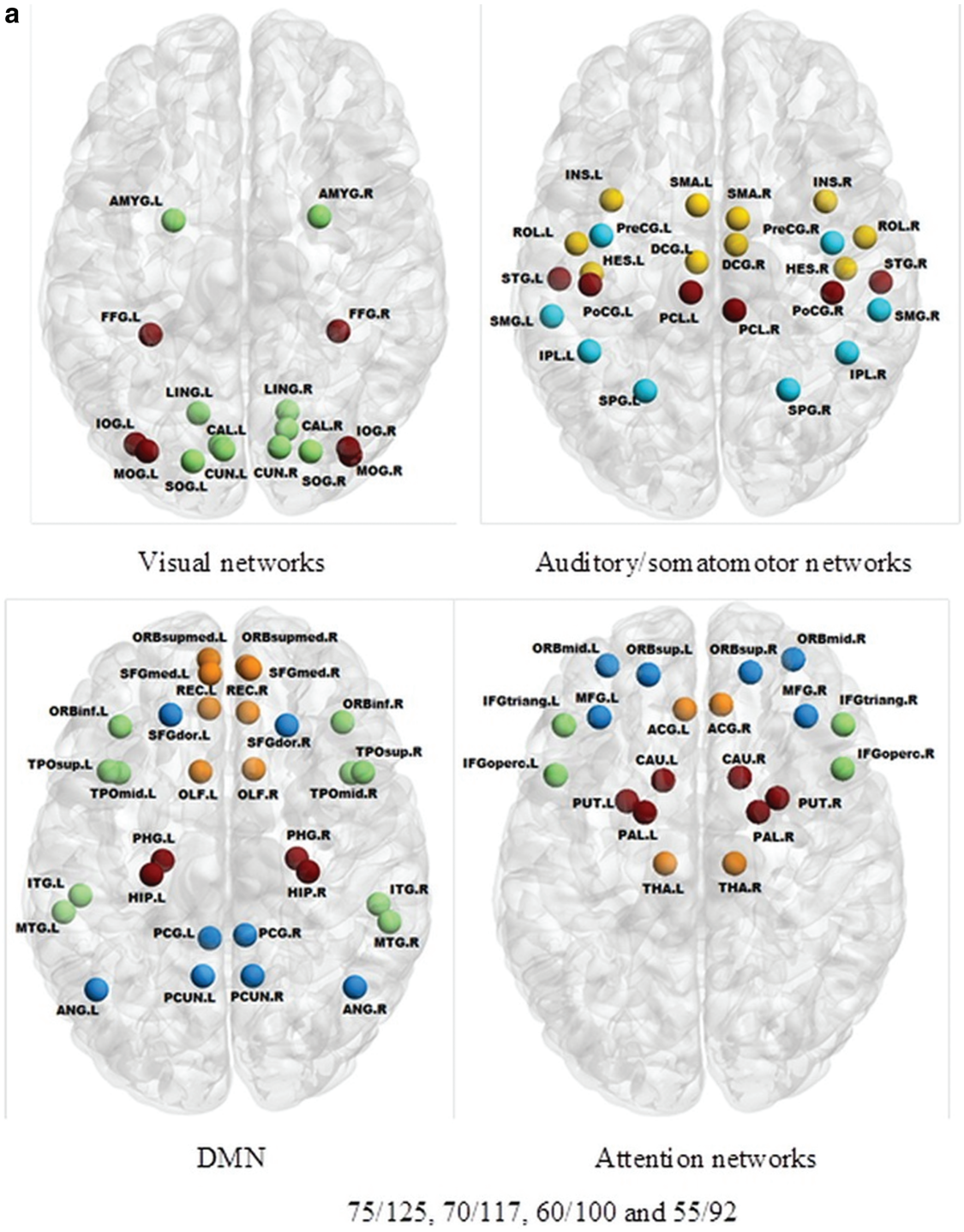

FIG. 5. Results of GNetHiClus for different subsample and sample sizes. Visual network, SMN, DMN, and attention network at the third level (Fig. 3) are displayed with nodes mapped onto the AAL template. In each of these four networks, the lowest level indivisible subnetworks extracted are represented in different colors, with each color corresponding to a green node in Figure 3. Clustering results are shown for the following subsampling/sample sizes: **(a)** 75/125, 70/117, 60/100, and 55/92; **(b)** 65/108; and **(c)** 50/83. For better appreciation of the differences of extracted subnetworks relative to that from 80/133 (i.e., the results considered as the ''ground truth''), such differences are indicated with circles of three different colors: blue, red, and black circles representing the missing brain regions, extra brain regions, and overclustered clusters (i.e., a cluster is inappropriately divided into a few clusters at higher levels), respectively. See Table 1 for the definition of node acronyms. AAL, automated anatomical labeling. Color images are available online. *Continued.*

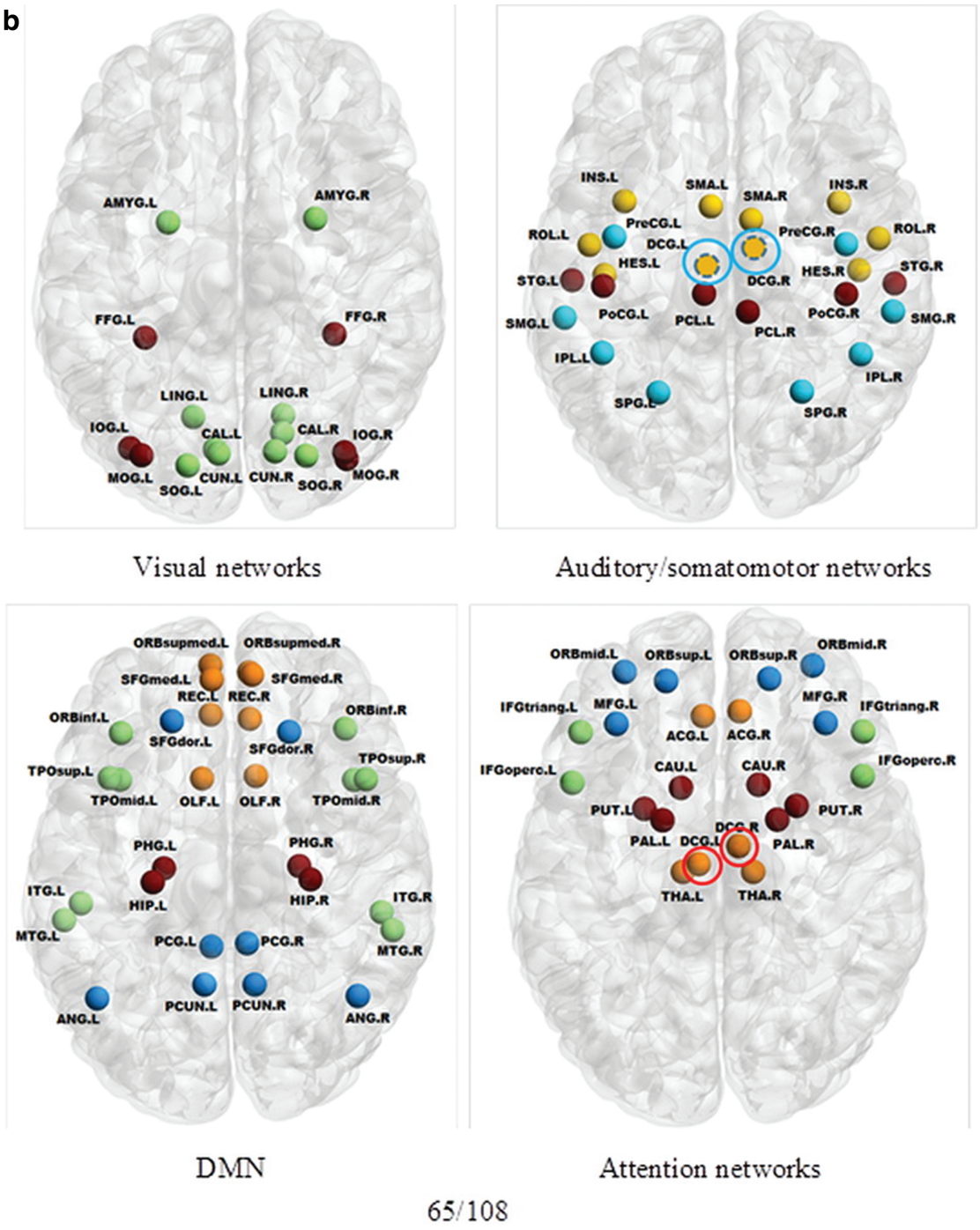

FIG. 5. (Continued).

network are commonly detected from resting-state human brain using either BOLD or arterial spin labeling perfusion fMRI (Beckmann et al., 2005; Liang et al., 2012). Similarly, with GNetHiClus, visual network, sensorimotor/auditory network, DMN, and attention network are extracted. Of note, due to the differences in the fundamental algorithms, there are a few apparent differences in the subnetwork extraction between previous findings and our clustering results: using ICA, the sensorimotor and the auditory network are identified as a single cluster (Beckmann et al., 2005; Liang et al., 2012); using GNetHiClus, the attention, the executive control, and the subcortical network are identified as a single cluster.

Note that we are not implying the proposed method is the only available approach to achieve hierarchical brain decomposition, but we described a novel approach to achieve this goal that has a number of important features.

Overall, the aim of this study is to recursively divide a large group of objects (i.e., brain regions in this study) into relatively small groups that share certain similarities given

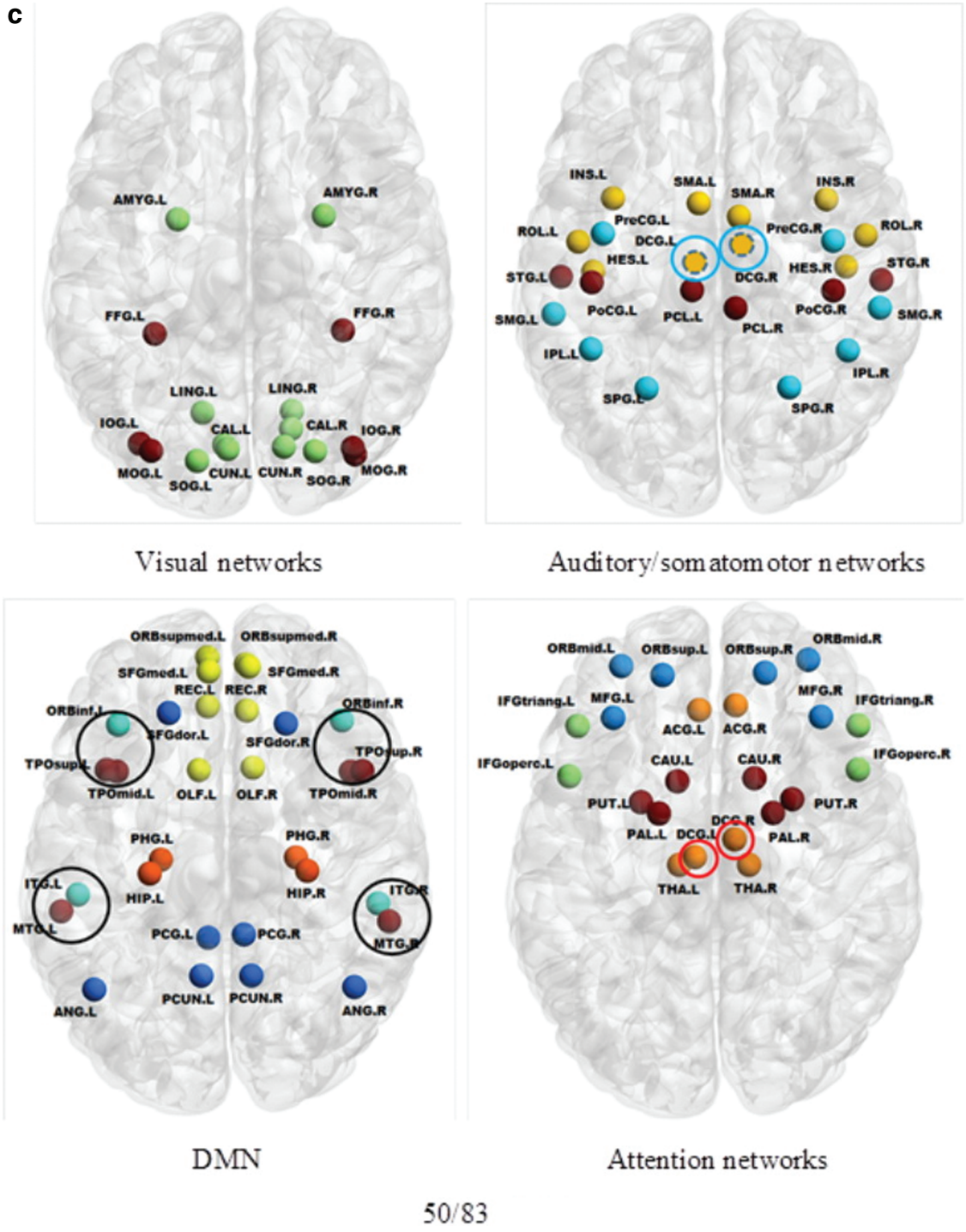

FIG. 5. (Continued).

the relationships (i.e., network connection strength) among all objects. This has been commonly perceived as community structure detection (or network modularity analysis), with which methods for analyzing network modularity have been developed (Newman and Girvan, 2004; Newman, 2006). These methods have been successfully employed for analyzing network modularity, but hierarchical network structure cannot be uncovered. Importantly, our method provides the capability of revealing hierarchical subnetworks across different levels, a goal that was also achieved in two previous studies (Ferrarini et al., 2009; Meunier et al., 2009b). However, direct comparisons between the proposed method and these two previous approaches were not conducted for the following reasons: (1) both previous studies focused on binary networks by setting arbitrary thresholds, instead of weighted networks in our study; (2) Ferrarini et al. (2009) employed partial correlations instead of full correlation as in our study.

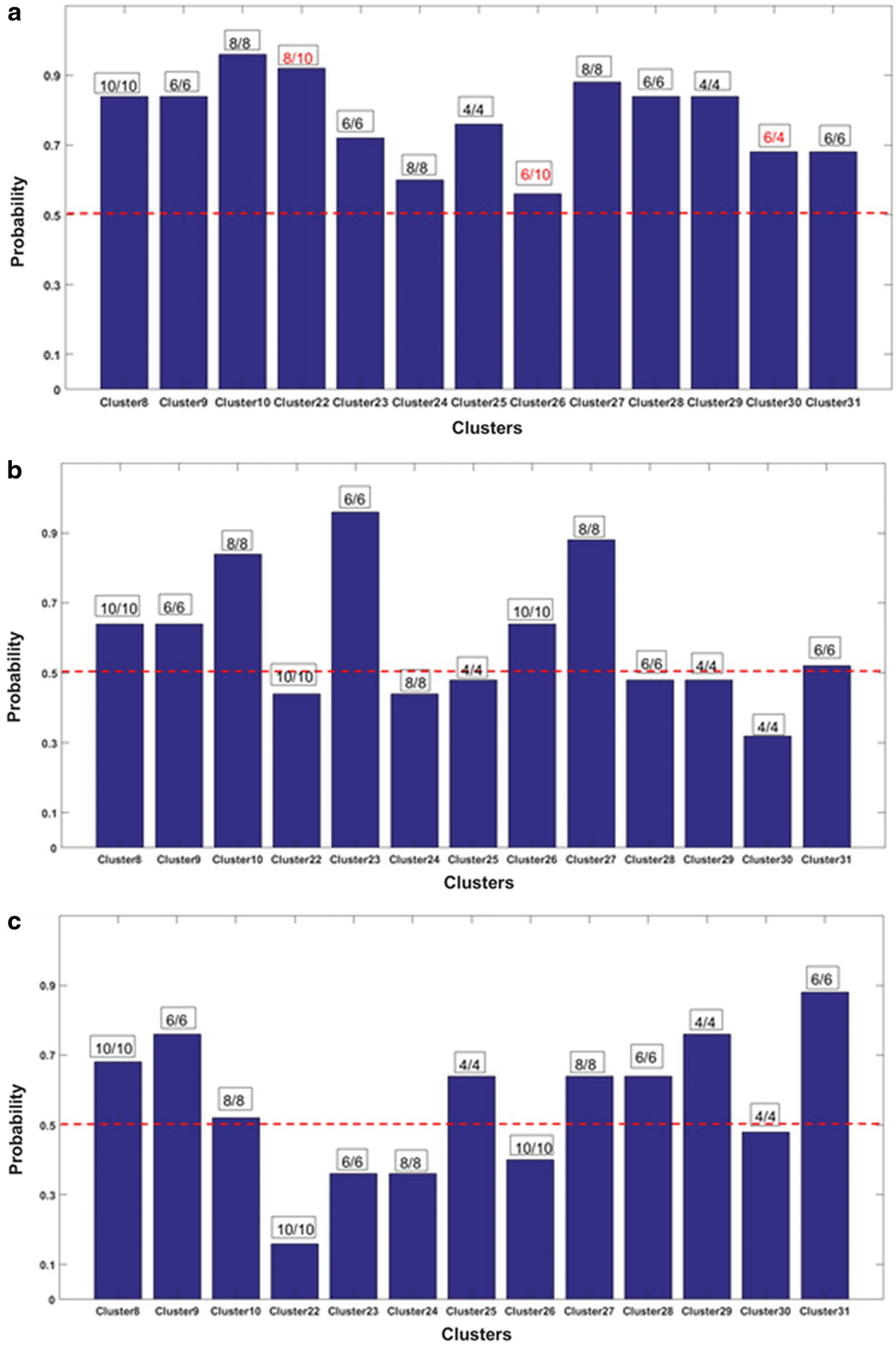

FIG. 6. Evaluation of the probability of each extracted subnetwork. The probability values of 13 subnetworks are shown for all six subsampling/sample sizes: **(a)** 50/83; **(b)** 55/92; **(c)** 60/100; **(d)** 65/108; **(e)** 70/117; and **(f)** 75/125. On top of each bar, a pair of numbers is shown, representing extracted number of regions/true number of regions; numbers of such clusters that are inconsistent with ground truth are shown in red. In addition, a red line corresponding to threshold = 0.5 is drawn to show extracted clusters with relatively low threshold (i.e., <0.5). Color images are available online. *Continued*.

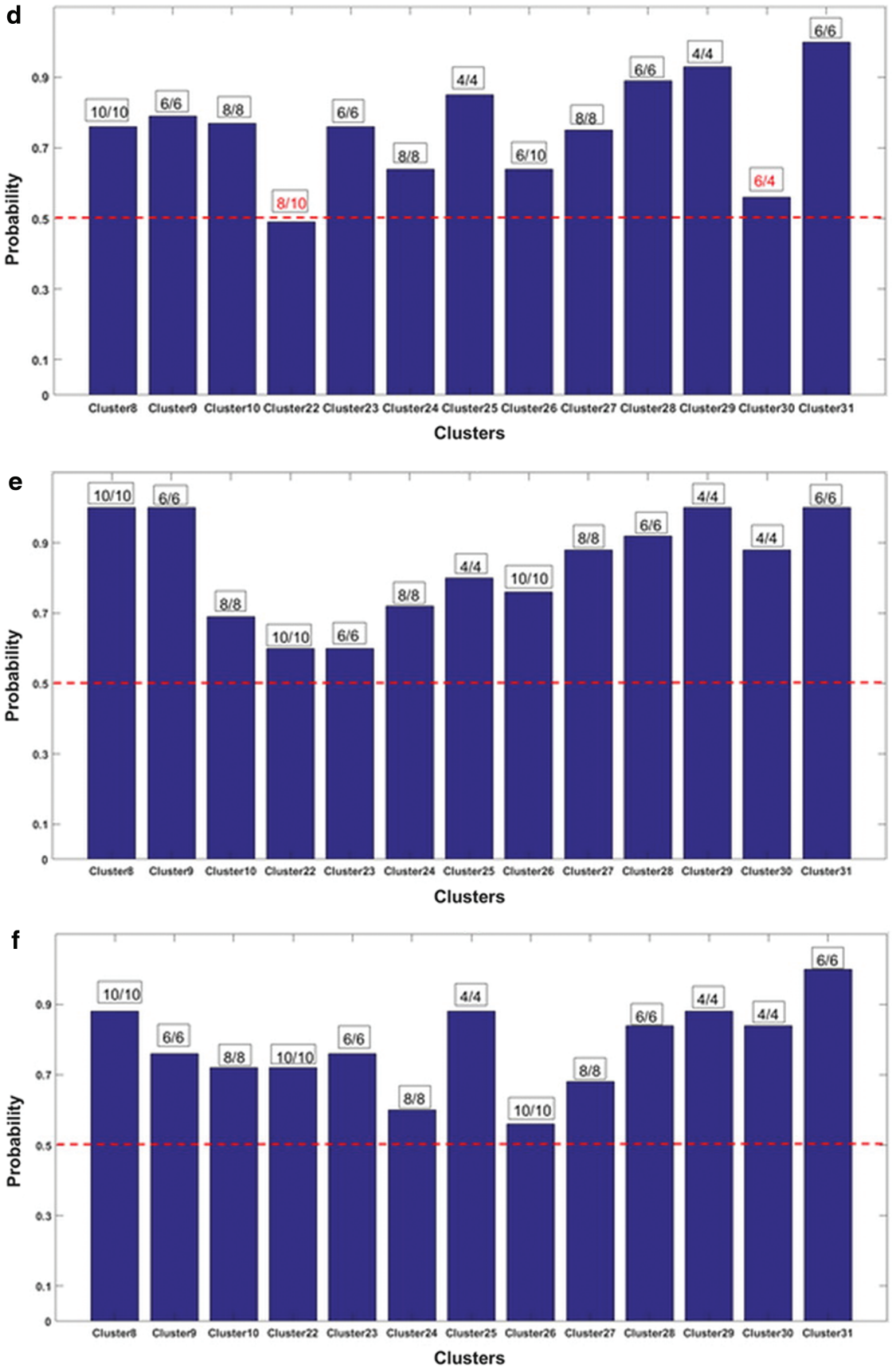

FIG. 6. (Continued).

Effect of sample size on group-level analysis

Given the use of bootstrapping approach, subsampling/sample size could be one of the most important factors that affect the final estimation of the clustering results. Our results have shown that reliable clustering results can be achieved with relatively high probability for subsampling/sample size of 70/117 and 75/125. Although fully consistent results cannot be obtained with lower number of subsampling/sample sizes, that is, 65/108, further inspection reveals that only one pair of brain regions is not desirably extracted into the right cluster. It should be noted that fully consistent clustering results have also been achieved for 60/100 and 55/92, although with a few clusters of low probability. Taken together, these clustering results suggest that the proposed method, GNetHiClus, can achieve satisfactory clustering outcomes (i.e., no more than a pair of brain regions is incorrectly assigned) with 75/125, 70/117, 65/108, 60/100, and 55/92. In contrast, regarding 50/83, four pairs of brain regions have not been correctly clustered into right clusters, indicating its inferior performance in comparison with those clustering results obtained with larger subsample sizes/sample sizes. Furthermore, among all extracted subnetworks, although visual subnetworks are the most reliable, auditory/somatomotor subnetworks are the least reliable, whose reliability cannot be consistently improved by increasing the subsampling/sampling sizes; further investigation is, therefore, warranted. Regardless, results have demonstrated that acceptable clustering results should still be achievable by using the proposed method, GNetHiClus, with reduced total number of subjects.

Use of the bilateral homotopic constraint

In GNetHiClus, homotopic brain regions are explicitly constrained to be clustered in the same subnetwork to minimize the noise effect (see also ''Limitations'' section). A previous study has shown that spontaneous brain activity is highly correlated across homotopic regions between left and right hemispheres (Tyszka et al., 2011), which suggests that homotopic regions tend to be cooperative with each other to achieve optimal brain function. Interestingly, such a bilateral symmetric property was also shown in both healthy adults and patients with complete agenesis of corpus callosum (Tyszka et al., 2011). This supports the plausibility of the bilateral homotopic constraint adopted in GNetHiClus.

Limitations

In this study, GNetHiClus is evaluated by using a relatively large number of subjects, based on the high-quality HCP MRI data sets. Our results have shown that a total number of $\sim$90 subjects are required to achieve acceptable clustering results, which might not be readily achievable for some clinical studies, due to the relatively small number of patients often collected. However, such a sample size should be feasible for multisite studies that collect relatively large cohorts, which is becoming increasingly common in the field.

The use of the bootstrapping technique in GNetHiClus imposes a computational demand, discouraging its application to high-spatial resolution parcellation schemes (Glasser et al., 2016)—or alternatively requiring the access to high-performance computing. We acknowledge that low-resolution AAL parcellation used in this study also has limitations. However, the parcellation was selected primarily based on the following reasons: (1) it is a commonly used parcellation template, (2) the empirical evidence of hierarchical modularity is likely to be more accessible, and (3) finer parcellation might be more informative, but it dramatically increases the computational burden.

For healthy subjects, the homotopic constraint (Tyszka et al., 2011) can be employed for achieving more reliable clustering results by removing uncertainty in assigning brain regions into specific clusters. However, it is not clear whether this constraint is still appropriate for patients with abnormal brain functions. To deal with such scenarios, further work is warranted to develop an alternative reliable approach for extracting subnetworks from group of patients with abnormal brain function without having to explicitly rely on a homotopic constraint. Future studies on the validity of the bilaterally homotopic constraint might provide important information on understanding the mechanism for maintaining normal brain function.

Conclusion

In this study, we proposed a novel approach, GNetHiClus, for hierarchically clustering the brain network into subnetworks, without relying on applying arbitrary thresholds to cross-coefficients. Importantly, rather than performing clustering at individual level, we employed a multisubject approach, sharing information across subjects. Our results show that GNetHiClus can hierarchically cluster brain functional network into specialized subnetworks that fulfill specialized tasks; conversely, information processed by specialized lower level subnetworks is integrated into upper level for achieving optimal efficiency. Our findings are consistent with the concept of network segregation and integration. This proposed technique should promote the understanding of brain network from a hierarchical point of view.

Acknowledgments

We are grateful to the National Health and Medical Research Council (NHMRC) of Australia, the Australian Research Council (ARC), and the Victorian Government's Operational Infrastructure Support Program for their support. Data were provided, in part, by the HCP, 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.

Author Disclosure Statement

No competing financial interests exist.

References

Achard S, Salvador R, Whitcher B, Suckling J, Bullmore E. 2006. A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. J Neurosci 26:63–72.

Bassett DS, Bullmore ET. 2017. Small-world brain networks revisited. Neuroscientist 23:499–516.

Bassett DS, Greenfield DL, Meyer-Lindenberg A, Weinberger DR, Moore SW, Bullmore ET. 2010. Efficient physical embedding of topologically complex information processing networks in brains and computer circuits. PLoS Comput Biol 6:e1000748.

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

Chen ZJ, He Y, Rosa-Neto P, Germann J, Evans AC. 2008. Revealing modular architecture of human brain structural networks by using cortical thickness from MRI. Cereb Cortex 18:2374–2381.

Chung F. 1997. Spectral graph theory. In Conference Board of the Mathematical Sciences, Washington.

Danon L, Diaz-Guilera A, Duch J, Arenas A. 2005. Comparing community structure identification. J Stat Mech-Theory Exp. DOI: 10.1088/1742-5468/2005/09/P09008.

Donath WE, Hoffman AJ. 1973. Lower bounds for partitioning of graphs. IBM J Res Dev 17:420–425.

Ferrarini L, Veer IM, Baerends E, Van Tol MJ, Renken RJ, Van Der Wee NJ, et al. 2009. Hierarchical functional modularity in the resting-state human brain. Hum Brain Mapp 30:2220–2231.

Fiedler M. 1973. Algebraic connectivity of graphs. Czech Math J 23:298–305.

Fred A, Jain A. Robust data clustering. In 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 18, 2003, Madison, WI.

Garrison KA, Scheinost D, Finn ES, Shen X, Constable RT. 2015. The (in)stability of functional brain network measures across thresholds. Neuroimage 118:651–661.

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

Kumar A, Rai P, Daume III, H. Co-regularized multi-view spectral clustering. In Proceedings of the 24th International Conference on Neural Information Processing Systems. Granada, Spain: 2011, pp. 1413–1421.

Lam L, Suen CY. 1997. Application of majority voting to pattern recognition: an analysis of its behavior and performance. IEEE Trans Syst Man Cybern A Syst Hum 27:553–568.

Liang X, Connelly A, Calamante F. 2014. Graph analysis of resting-state ASL perfusion MRI data: nonlinear correlations among CBF and network metrics. Neuroimage 87:265–275.

Liang X, Connelly A, Calamante F. 2016. A novel joint sparse partial correlation method for estimating group functional networks. Hum Brain Mapp 37:1162–1177.

Liang X, Tournier JD, Masterton R, Connelly A, Calamante F. 2012. A k-space sharing 3D Grase pseudocontinuous ASL method for whole-brain resting-state functional connectivity. Int J Imaging Syst Technol 22:37–43.

Liang X, Vaughan DN, Connelly A, Calamante F. 2018. A novel group-fused sparse partial correlation method for simultaneous estimation of functional networks in group comparison studies. Brain Topogr 31:364–379.

Meunier D, Achard S, Morcom A, Bullmore E. 2009a. Age-related changes in modular organization of human brain functional networks. Neuroimage 44:715–723.

Meunier D, Lambiotte R, Fornito A, Ersche KD, Bullmore ET. 2009b. Hierarchical modularity in human brain functional networks. Front Neuroinform 3:37.

Moretti P, Munoz MA. 2013. Griffiths phases and the stretching of criticality in brain networks. Nat Commun 4:2521.

Mottron L, Dawson M, Soulieres I, Hubert B, Burack J. 2006. Enhanced perceptual functioning in autism: an update, and eight principles of autistic perception. J Autism Dev Disord 36:27–43.

Newman MEJ. 2006. Modularity and community structure in networks. Proc Natl Acad Sci U S A 103:8577–8582.

Newman MEJ, Girvan M. 2004. Finding and evaluating community structure in networks. Phys Rev E 69:026113.

Ng AY, Jordan MI, Weiss Y. On spectral clustering: analysis and an algorithm. In Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic. Vancouver, Canada: 2002, pp. 849–856.

Ng B, Varoquaux G, Poline JB, Thirion B, 2013. A novel sparse group Gaussian graphical model for functional connectivity estimation. Inf Process Med Imaging 23:256–267.

Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, et al. 2011. Functional network organization of the human brain. Neuron 72:665–678.

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.

Shang J, Lui S, Meng Y, Zhu H, Qiu C, Gong Q, et al. 2014. Alterations in low-level perceptual networks related to clinical severity in PTSD after an earthquake: a resting-state fMRI study. PLoS One 9:e96834.

Simon H. 1962. The architecture of complexity. Proc Am Philos Soc 106, 467–482.

Tyszka JM, Kennedy DP, Adolphs R, Paul LK. 2011. Intact bilateral resting-state networks in the absence of the corpus callosum. J Neurosci 31:15154–15162.

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. 2002. Automated anatomical labeling of activations in Spm using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15:273–289.

Van Essen DC, Smith SM, Barch DM, Behrens TE, Yacoub E, Ugurbil K, et al. 2013. The WU-Minn human connectome project: an overview. Neuroimage 80:62–79.

Varoquaux G, Gramfort A, Poline JB, Thirion B. 2010. Brain covariance selection: better individual functional models using population prior. Adv Neural Inf Process Syst 23:2334–2342.

Address correspondence to:
Xiaoyun Liang
The Florey Institute of Neuroscience and Mental Health
Melbourne Brain Centre
245 Burgundy Street
Heidelberg 3084
Victoria
Australia

E-mail: imagetechliang@gmail.com