

Atomic connectomics signatures for characterization and differentiation of mild cognitive impairment

Jinli Ou · Li Xie · Xiang Li · Dajiang Zhu · Douglas P. Terry · A. Nicholas Puent · Rongxin Jiang · Yaowu Chen · Lihong Wang · Dinggang Shen · Jing Zhang · L. Stephen Miller · Tianming Liu

© Springer Science+Business Media New York 2014

Abstract In recent years, functional connectomics signatures have been shown to be a very valuable tool in characterizing and differentiating brain disorders from normal controls. However, if the functional connectivity alterations in a brain disease are localized within sub-networks of a connectome, then accurate identification of such disease-specific sub-networks is critical and this capability entails both fine-granularity definition of connectome nodes and effective clustering of connectome nodes into disease-specific and non-disease-specific sub-networks. In this work, we adopted the recently developed DICCCOL (dense individualized and common connectivity-based cortical landmarks) system as a fine-granularity high-resolution connectome construction

method to deal with the first issue, and employed an effective variant of non-negative matrix factorization (NMF) method to pinpoint disease-specific sub-networks, which we called atomic connectomics signatures in this work. We have implemented and applied this novel framework to two mild cognitive impairment (MCI) datasets from two different research centers, and our experimental results demonstrated that the derived atomic connectomics signatures can effectively characterize and differentiate MCI patients from their normal controls. In general, our work contributed a novel computational framework for deriving descriptive and distinctive atomic connectomics signatures in brain disorders.

Keywords Resting state fMRI · Brain networks · Functional connectome · MCI · NMF · DICCCOL

J. Ou · L. Xie · R. Jiang · Y. Chen
School of Biomedical Engineering & Instrument Science, Zhejiang University, Hangzhou, China

X. Li · D. Zhu · T. Liu (✉)
Cortical Architecture Imaging and Discovery Lab, Department of Computer Science and Bioimaging Research Center, The University of Georgia, Athens, GA, USA
e-mail: tlju@cs.uga.edu

D. P. Terry · A. N. Puent · L. S. Miller
Department of Psychology, The University of Georgia, Athens, GA, USA

D. Zhu · L. S. Miller
Bioimaging Research Center, The University of Georgia, Athens, GA, USA

J. Zhang
Department of Statistics, Yale University, New Haven, CT, USA

D. Shen
Department of Radiology, UNC, Chapel Hill, NC, USA

L. Wang
Department of Biomedical Engineering, Tsinghua University, Beijing, China

Introduction

Neuroimaging has been an important technique in understanding the structural and functional alterations in brains with various psychiatric/mental disorders. Recently, functional connectomes constructed from neuroimaging data have emerged as a powerful tool in characterizing and differentiating brain disorders (Camchong et al. 2011; Cocchi et al. 2012; Arbabshirani et al. 2013; Fornito and Bullmore 2014; Li et al. 2014; Zhu et al. 2014; Li et al. 2013; Venkataraman et al. 2012). For instance, functional connectomics signatures were derived for characterization and differentiation of mild cognitive impairment (MCI) (Wee et al. 2012; Zhu et al. 2014), Alzheimer's disease (AD) (Wang et al. 2007; Greicius et al. 2004; Supekar et al. 2008), posttraumatic stress disorder (Li et al. 2014; Lanius et al. 2005; Gilboa et al. 2004) and prenatal cocaine exposure affected brains (Li et al. 2013; Santhanam et al. 2011). In general, those functional connectomics signatures derived in those previous studies were typically over

whole-brain networks, i.e., the whole networks of a connectome. If the functional alterations in a brain condition are localized within specific functional sub-networks of a connectome, whole-brain functional connectomes may not be effective and sensitive enough to characterize and differentiate the brain condition. Therefore, functional connectomics signatures within sub-networks need to be explored, instead.

From a technical perspective, the major barrier to brain connectivity analysis is the lack of large-scale common and consistent brain landmarks as a generic brain reference system, which provides network nodes for constructing connectivities within individual brains and for comparing connectivities across different brains. Due to the lack of generic and large-scale brain landmarks, fine-granularity connectivity networks of those functional connectomics signatures derived in previous studies (Wee et al. 2012; Wang et al. 2007; Greicius et al. 2004; Supekar et al. 2008; Lanius et al. 2005; Gilboa et al. 2004; Santhanam et al. 2011) are still largely unknown. Thus, determination of common and consistent brain landmarks across different brains is perhaps one of the greatest challenges in human brain mapping (Derrfuss and Mar 2009; Poldrack 2012; Liu 2011). Recently, a large set of reliable and dense cortical landmarks has been identified and publicly released (<http://dicccol.cs.uga.edu>), named Dense Individualized and Common Connectivity-based Cortical Landmarks (DICCCOL) (Zhu et al. 2013). The neuroscience basis is that each cytoarchitectonic area possesses a unique set of extrinsic inputs and outputs, namely the “connectional fingerprint”, which largely determine the functions that the area can perform (Passingham et al. 2002). In the current stage, the DICCCOL system has identified 358 consistent cortical ROIs, the locations of which were optimized via maximizing the group-wise consistency of DTI-derived fiber shape connectivity patterns (Zhu et al. 2013). Importantly, these 358 DICCCOLs could be accurately predicted and localized in an individual brain based only on DTI data (Zhu et al. 2013). Thus, in this work, we employed these DICCCOL landmarks as network nodes for fine-granularity whole-brain connectome construction. Because of the intrinsically structural and functional correspondences of the DICCCOL system across brains, functional connectomes constructed from different brains can be readily pooled and integrated. The 358 DICCCOL landmarks on the reconstructed WM/GM cortical surfaces were illustrated in Fig. 1. Notably, it should be pointed out that the current 358 DICCCOLs are not the whole-cortex parcellation, but are sampled consistent ROIs.

Generally speaking, the human connectome (Kennedy 2010) is a complex network with highly functional segregation and integration (Bassett and Bullmore 2006; Stam 2010; Sporns et al. 2004; Sporns 2011, 2013), and can be functionally divided into many sub-networks, for instance, the default mode, working memory, vision, auditory, attention and

emotion, some of which may overlap each other. The human connectome can be characterized by a large-scale functional connectivity matrix, e.g., denoted by \mathbf{X} . From a mathematical perspective, any given matrix \mathbf{X} can be factorized as a linear combination of some orthogonal bases. Because of the non-negative constraints allowing only additive, not subtractive, combinations, non-negative matrix factorization (NMF) can effectively learn part-based representations (Lee and Seung 1999; Yang et al. 2007) of a composite object.

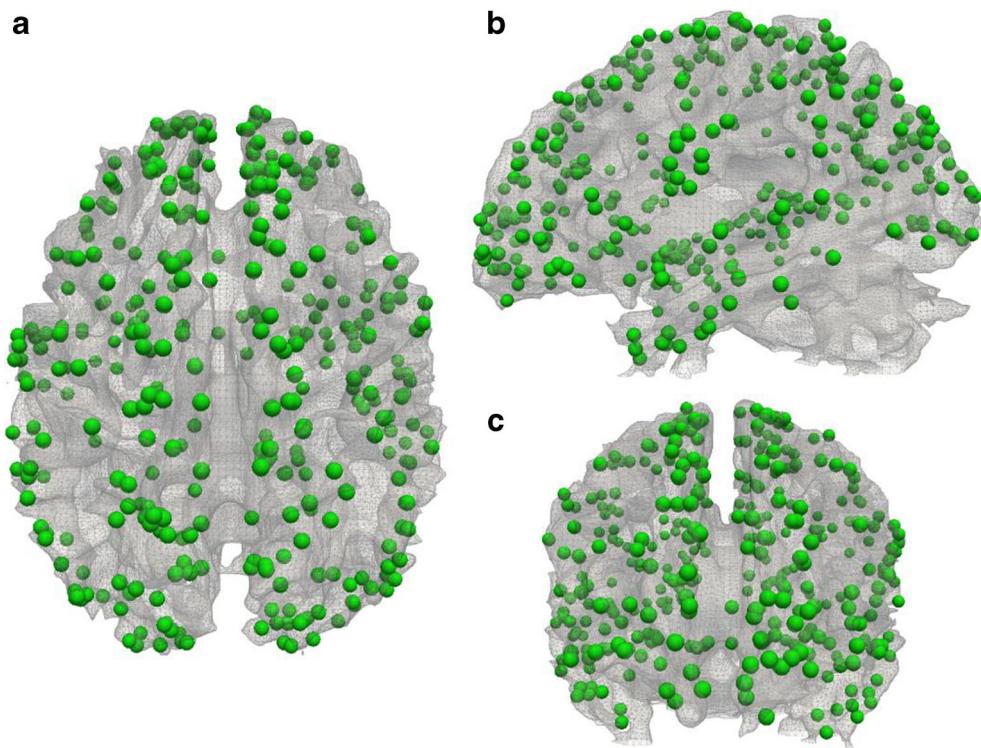
Motivated by the composite structure of the functional connectome and the success of NMF in learning part-based representations of a composite object, in our previous work (Ou et al. 2014), we applied an effective variant of NMF, i.e., the divergence-based projective non-negative matrix factorization (D-PNMF) (Yang and Oja 2010), along with a Bayesian change point model, to investigate the atomic dynamic functional interaction patterns in attention-deficit/hyperactivity disorder. To achieve a more general and feasible computational framework, in this work, we first integrated two MCI datasets from two different research centers and then apply the D-PNMF approaches directly to pooled datasets, for investigating the connectivity patterns within sub-networks of a connectome, which we called atomic functional connectomes (AFCs). As a result, four AFCs were obtained, two from the MCI group and another two from the normal control (NC) group. Experimental results suggest that the four AFCs can be categorized into distinctive and common pairs, which indicate the connectivity patterns within disease-specific and non-disease-specific sub-networks. The distinctive pair shows significantly increased and decreased connectivity patterns involved in MCI, as compared with NC. In addition, the four AFCs also have high classification accuracy with both good sensitivity and specificity, and can be viewed as “connectomics signatures” for MCI and NC. Essentially, the work in this paper implicates that MCI could be a disorder of connectivity between components of a connectome.

Materials and methods

Overview

The pipeline of the proposed computational framework is summarized in Fig. 2. First, after the fMRI/DTI data preprocessing (step 1), the 358 consistent DICCCOL landmarks are predicted and localized in the DTI data of each brain (step 2). Then, the fMRI data is co-registered into the DTI space using FSL FLIRT (step 3). The fMRI BOLD signals are then extracted from the co-registered fMRI data by averaging in a small neighborhood (3 mm radius) for each DICCCOL (Zhu et al. 2013) (step 4). Subsequently, a whole-brain functional connectome is evaluated with Pearson correlations between every pair of fMRI BOLD signals and characterized by a

Fig. 1 Visualization of the 358 DICCCOL ROIs on the cortical surfaces in the top **a**, lateral **b** and rear **c** views. The ROIs are shown as green spheres. The visualizations of the DTI-derived fiber patterns of these ROIs are referred to: <http://dicccol.cs.uga.edu>



symmetric connectivity matrix (step 5). The D-PNMF approaches are then applied to the pooled connectivity matrices for learning AFCs (step 6). Finally, we compare the approximation residual errors of each FCV from its reconstructive FCVs based on the learned AFCs for classification purpose (step 7). Details of these steps are described in detail in the following sections.

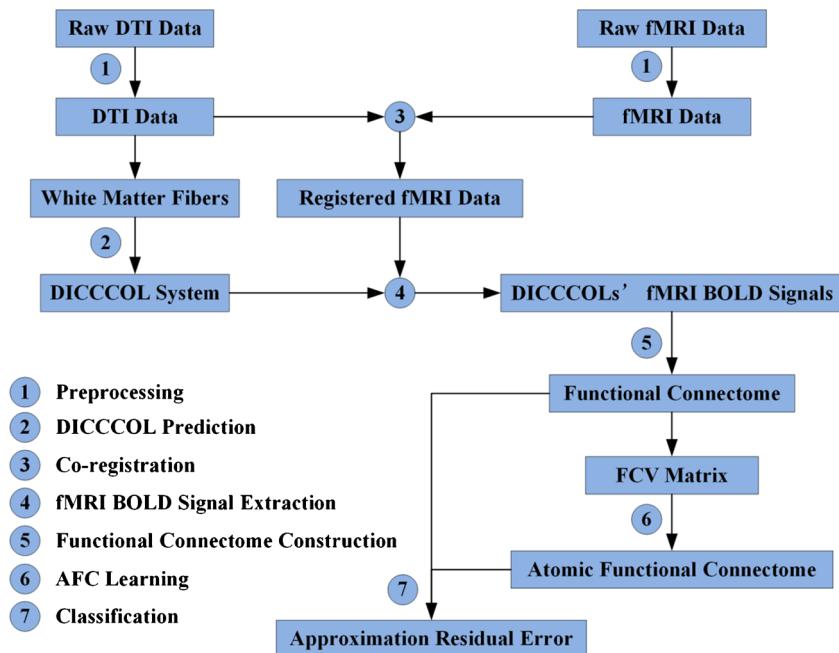
Subjects

To infer and investigate connectomics signatures in MCI, two independent MCI datasets from two different research centers (Duke Medical Center and UGA Bioimaging Research Center) were used in this work.

Dataset 1 included 27 participants (10 MCI patients and 17 normal controls) who were recruited and scanned in the Duke-UNC Brain Imaging and Analysis Center (BIAC). MCI patients were diagnosed according to NACC procedures and NINCDS-ADRDA diagnostic guidelines, and their diagnoses were based upon available data from a general neurological examination, neuropsychological assessment evaluation, collateral and subject symptom and functional capacity reports. Confirmation of diagnosis for all subjects was made via expert consensus panels at the Joseph and Kathleen Bryan Alzheimer's Disease Research Center (Bryan ADRC) and the Department of Psychiatry at Duke University Medical Center. The inclusion and exclusion criteria for dataset 1 can be found in Wee et al. (2012). To be self-contained and brief,

the inclusion criteria for MCI patients are: 1) age >55 years and any race; 2) recent worsening of cognition, but still functioning independently; 3) Mini Mental State Examination (MMSE) score between 24 and 30; 4a.) score ≤ -1.5 SD on at least two Bryan ADRC cognitive battery memory tests for single-domain amnestic MCI; or 4b.) score ≤ -1.5 SD on at least one of the formal memory tests and score ≤ -1.5 SD on at least one other cognitive domain task for multi-domain MCI; 5) score 4 or lower for baseline Hachinski; 6) no psychological symptoms or history of depression; 7) exhibit no dementia according to the NINCDS-ADRDA or DSM-IV-TR criteria; and 8) ability to give informed consent and follow study protocols. While the inclusion criteria for NC subjects are: 1) age >55 years and any race; 2) adequate visual and auditory acuity to finish neuropsychological testing; 3) a normal non-focal neurological examination score; 4) no self-report of neurological or depressive illness; 5) exhibits no depression according to the Diagnostic Interview Schedule portion of the Duke Depression Evaluation Schedule; 6) a score >-1 SD on any formal memory tests and a score >-1 SD any formal executive function or other cognitive test; and 7) ability to give informed consent and follow study protocols. Subjects were excluded, who met the following criteria: 1) any traditional MRI contraindication, such as foreign metallic implants or pacemakers; 2) any of other Axis I psychiatric disorders; 3) any physical or intellectual disability preventing assessments; 4) any head injury or neurological disorder associated with MRI abnormalities, such as dementia, brain tumors, epilepsy,

Fig. 2 The pipeline of the computational framework is composed of seven steps: 1) preprocessing of the DTI and fMRI data; 2) DICCCOL prediction in the DTI data; 3) co-registration of the fMRI data into the DTI space; 4) fMRI BOLD signal extraction by averaging in a small neighborhood (3 mm radius) for each predicted DICCCOL; 5) functional connectome construction from the fMRI BOLD signals; 6) AFC learning via the D-PNMF approaches; 7) classification based on the approximation residual errors



demyelinating diseases and so on; and 5) any prescription medication (or nonprescription drugs) with known neurological effects.

Dataset 2 included 24 participants (12 MCI patients and 12 normal controls) who were recruited from a 10-county area around Athens-Clarke County, Georgia. Initial recruitment was through advertisements and community contacts, which was developed by the UGA Neuropsychology and Memory Assessment Laboratory. Informed consent was obtained after reviewing detailed written information about the experimental protocol, which was approved by the UGA IRB. Participants completed comprehensive evaluations including MRI compatibility screening, participant and collateral interviews regarding relevant social and medical history by a trained interviewer certified in dementia rating, self, and collateral reports of activities of daily living (ADLs), and neuropsychological testing. The clinical dementia rating (CDR) staging decisions (group placement) based on CDR guidelines (Hughes et al. 1982) was made upon the information from the initial interview. The inclusion criteria for dataset 2 can be also found in Faraco et al. (2013) and Puente et al. (2014). Briefly, the inclusion criteria for dataset 2 are: 1) age between 65 and 85 years; 2) a reliable collateral, literate, and no self-reported history of a neurological disorder; 3) capacity to give informed consent and follow the entire MR scanning protocol; and 4) compatibility with the MRI environment. For MCI patients, there were another two extra inclusion criteria: 1) self and collateral report of memory decline, and 2) objective memory impairment within the memory domain of the CDR. Participants were categorized as NC if they received a global CDR score of 0, as MCI if they received a global CDR score of 0.5, and as dementing if they received a global CDR \geq 1.

Demographics of the participants in two Datasets are summarized in Table 1.

Data acquisition and preprocessing

Dataset 1: The DTI and fMRI data were acquired on a 3 T GE MRI scanner at the Duke-UNC BIAC, under Duke IRB approval and with the following parameters. For fMRI, the parameters were: matrix=64 \times 64, slice thickness=3.8 mm, field of view (FOV)=256 \times 256 mm 2 , TR (repetition time)=2 s, TE (echo time)=32 ms. For DTI, 25 direction diffusion-weighted whole-brain volumes were acquired axially parallel to the AC-PC using the following parameters: b=0,1000 s/mm 2 , TR=17 s, TE=78 ms, matrix=128 \times 128, FOV=256 \times 256 mm 2 , resulting in an imaging resolution of 2 \times 2 \times 2 mm 3 .

Dataset 2: The DTI and fMRI data were scanned on a 3 T GE Signa HDx MRI system at the UGA, under UGA IRB approval. For fMRI, the parameters were: matrix=64 \times 64, slice thickness=4 mm, FOV=256 \times 256 mm 2 , TR=5 s, TE=25 ms. For DTI, 30 direction diffusion-weighted whole-brain volumes were acquired axially parallel to the AC-PC with the following parameters: b=0,1000 s/mm 2 , TR=17 s, TE=min full (minimum echo time needed, here the “min full” value for TE is 86 ms), matrix=128 \times 128, FOV=256 \times 256 mm 2 , resulting in an imaging resolution of 2 mm isotropic.

Briefly, the data preprocessing steps of the fMRI data included brain skull removal, motion correction, spatial smoothing, temporal prewhitening, slice time correction, global drift removal, and band pass filtering (0.01Hz~0.1Hz) (Li et al. 2012; Zhu et al. 2012). As the spontaneous BOLD signals fluctuate in a low frequency in resting state, mainly between 0.01 Hz and 0.1 Hz (Fransson 2005; Fox and

Table 1 Demographic information of the participants in Dataset 1 and Dataset 2

	Dataset 1		Dataset 2	
	NC	MCI	NC	MCI
No. of subjects	17	10	12	12
No. of males	9	5	2	5
Age (mean±SD)	74.2±8.6	72.1±8.2	72.3±5.1	78.1±4.8
Education (mean±SD)	17.7±4.2	16.3±2.4	16.1±2.6	14.7±3.6
MMSE (mean±SD)	28.4±1.5	29.4±0.9	28.3±1.7	25.5±2.5
Memory/Cognitive test score	>-1	≤-1.5	/	/
CDR score	/	/	0	0.5

Raichle 2007). This is the typical band that in resting state networks are of interest. It should be noted that the first few volumes have been removed before preprocessing. After removing the first few volumes, there were 150 volumes left in dataset 1 and 108 volumes left in dataset 2. Preprocessing steps of the DTI data included brain skull removal, motion correction, and eddy current correction. More details of the fMRI/DTI data preprocessing can be found in our previous publications (Li et al. 2012; Zhu et al. 2012, 2013).

After the fMRI/DTI data preprocessing, the 358 DICCCOL landmarks were predicted and localized in the DTI data of each individual brain via a data-driven strategy (Zhu et al. 2013). Briefly, the prediction process consists of three major steps: initial landmarks selection, optimization of landmark locations, and determination of group-wise consistent DICCCOLs (Zhu et al. 2013). More details of the DICCCOL prediction can be found in our previous publica-

tion (Zhu et al. 2013) and the publicly accessible DICCCOL website (<http://dicccol.cs.uga.edu>).

Functional connectome construction and vectorization

Because fMRI and DTI sequences are both EPI (echo planar imaging) sequences, their distortions tend to be similar and the misalignment between fMRI and DTI images is much less than that between fMRI and T1 images (Li et al. 2012). Thus, after DICCCOL prediction (Zhu et al. 2013) in the DTI data of each brain (Fig. 3a), the fMRI images were co-registered into the DTI space using FSL FLIRT. Then, the fMRI BOLD signals were acquired by averaging in a small neighborhood (3 mm radius) for each DICCCOL (Fig. 3b). Here, the pairwise functional connectivity was estimated as the absolute value of the Pearson correlations between every pair of fMRI BOLD signals (Fig. 3c), as follows:

$$R_{i,j} = \text{abs}(\text{corr}(S_i, S_j)), \text{ if } i \neq j; R_{i,j} = 0, \text{ if } i = j; FC = \left\{ R_{i,j} \mid i, j \in (1, 358) \right\} \quad (1)$$

where S_i denotes the fMRI BOLD signals extracted for the i -th DICCCOL, $R_{i,j}$ denotes the absolute value of the Pearson correlation between the fMRI BOLD signals extracted for the i -th and j -th DICCCOLs and FC denotes the functional connectome over whole-brain networks, which is characterized by a symmetrical non-negative 358×358 connectivity matrix. Considering the symmetry of the functional connectivity matrix, its upper triangular elements are concatenated to make a vector of $M=358 \times 357/2=63903$ features (Fig. 3d), denoted as functional connectome vector (FCV). Due to the intrinsically-established correspondences of DICCCOLs across individual brains, FCVs constructed from different brains can be readily pooled and integrated to make a single FCV matrix for further analysis (Fig. 3e).

Atomic functional connectome (AFC) learning

In our previous works (Yuan et al. 2013; Zhu et al. 2013), the functional roles of 358 DICCCOLs were annotated via meta-

analysis with the existing fMRI studies reported and aggregated in the BrainMap database. In total, 339 DICCCOLs have been functionally labeled with 55 functional sub-networks (Yuan et al. 2013) via the meta-analysis. Hence, the whole-brain functional connectome based on DICCCOLs can be viewed as a combination of connectivity patterns within different functional sub-networks, which we called atomic functional connectomes (AFCs). Based on the above methods in the section of “functional connectome construction and vectorization”, a functional connectome can be represented by a vector, i.e., FCV, and a non-negative FCV matrix can be obtained from a group. From a mathematical perspective, the AFCs can be viewed as bases and be resolved by a matrix factorization model (Trefethen and Bau III 1997), as follows.

$$\mathbf{X}_{M \times N} \approx \mathbf{W}_{M \times r} \mathbf{H}_{r \times N} \quad (2)$$

where columns of \mathbf{X} are the FCVs constructed for each individual, and columns of \mathbf{W} denote the bases, i.e., the

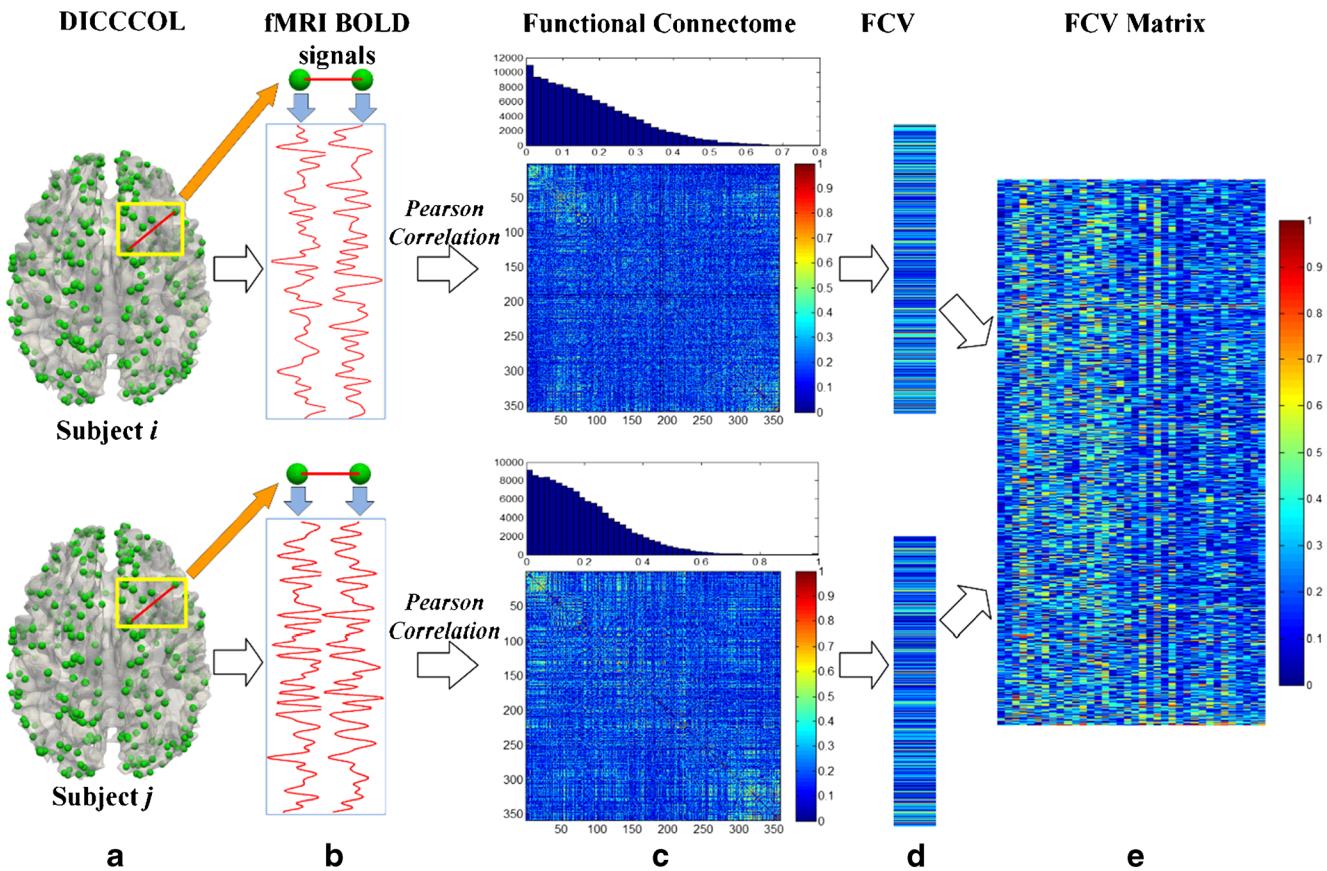


Fig. 3 Functional connectome construction and vectorization. **a** The predicted 358 DICCCOL landmarks on the cortical surfaces, shown as green spheres. **b** The fMRI BOLD signals extracted for two randomly selected DICCCOLs. **c** The functional connectivity matrices estimated from the fMRI BOLD signals with their corresponding histograms

showing the distribution of the values on the top of each of the connectivity matrices. **d** The vectorized representations of the functional connectivity matrices, i.e., FCVs. **e** The pooled FCV matrix composed of FCVs constructed from different brains. The vectors and matrices are color coded by using a heat color map, from dark blue (minimum) to dark red (maximum)

vectorized representatives of AFCs, which can be used to reconstruct \mathbf{X} with the coefficients \mathbf{H} . M denotes the dimension of FCV and AFC, and N denotes the number of subjects. The rank r denotes the rank of \mathbf{W} , which is also the number of AFCs. Here, we adopted the divergence-based projective non-negative matrix factorization (D-PNMF) (Yang and Oja 2010) approaches to learn the bases, i.e., AFCs. An example of AFC learning is illustrated in Fig. 4.

In this work, to reduce the parameters during NMF, \mathbf{H} is assumed to be the projection of \mathbf{X} onto \mathbf{W} , i.e., $\mathbf{H} = \mathbf{W}^T \mathbf{X}$. Thus, under the non-negative constraints on the elements of \mathbf{W} and \mathbf{H} , AFCs become the solution to the optimization problem:

$$\begin{aligned} \min_{\mathbf{W} \geq 0} \|\mathbf{X} - \mathbf{WW}^T \mathbf{X}\| &\Leftrightarrow \min_{\mathbf{W} \geq 0} D(\mathbf{X} \parallel \mathbf{Y}) \\ &= \min_{\mathbf{W} \geq 0} \sum_{i,j} \left(X_{i,j} \log \frac{X_{i,j}}{Y_{i,j}} - X_{i,j} + Y_{i,j} \right) \quad (\mathbf{Y} = \mathbf{WW}^T \mathbf{X}), \end{aligned} \quad (3)$$

where $\|\cdot\|$ denotes a matrix norm, and $D(\mathbf{X} \parallel \mathbf{Y})$ is the matrix divergence of \mathbf{X} from \mathbf{Y} , which reduces to the Kullback–Leibler divergence, when $\sum_{i,j} X_{i,j} = \sum_{i,j} Y_{i,j} = 1$. With the

gradient descent approach, the additive update rule for Eq. (3) becomes:

$$W_{i,j} \leftarrow W_{i,j} - \eta_{i,j} \frac{\partial D(\mathbf{X} \parallel \mathbf{WW}^T \mathbf{X})}{\partial W_{i,j}}, \quad (4)$$

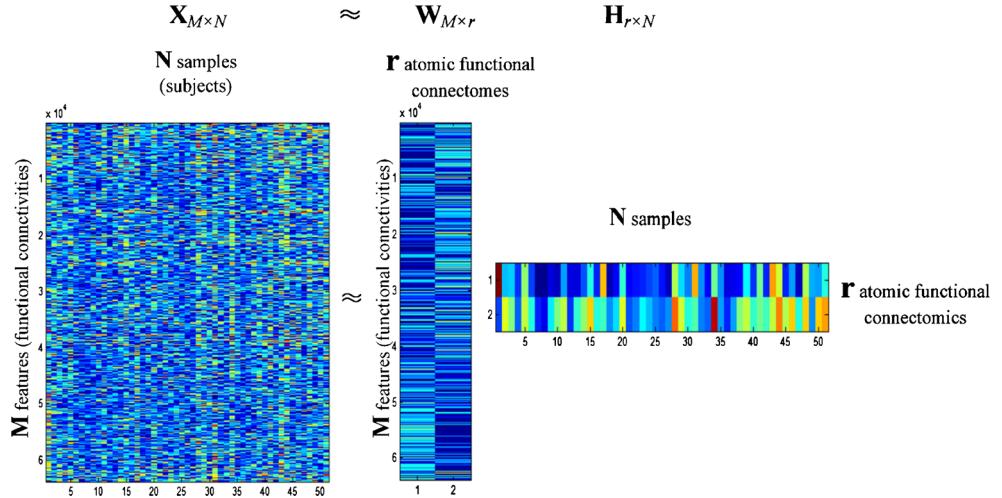
where $\eta_{i,j}$ is a non-negative step size. For the divergence, the gradient in Eq. (4) is:

$$\begin{aligned} \frac{\partial D(\mathbf{X} \parallel \mathbf{WW}^T \mathbf{X})}{\partial W_{i,j}} &= \sum_k \left((\mathbf{W}^T \mathbf{X})_{j,k} + \sum_l W_{l,j} X_{i,k} \right) - \\ &\quad \sum_k \frac{X_{i,k} (\mathbf{W}^T \mathbf{X})_{j,k}}{(\mathbf{WW}^T \mathbf{X})_{i,k}} - \sum_k X_{i,k} \sum_l \frac{W_{l,j} X_{l,k}}{(\mathbf{WW}^T \mathbf{X})_{l,k}}. \end{aligned} \quad (5)$$

The non-negative constraint on \mathbf{W} is guaranteed by positive initialization of \mathbf{W} and setting the step size in Eq. (4) as follows:

$$\eta_{i,j} = \frac{W_{i,j}}{\sum_k \left((\mathbf{W}^T \mathbf{X})_{j,k} + \sum_l W_{l,j} X_{i,k} \right)}. \quad (6)$$

Fig. 4 A rank-2 (r) matrix factorization on a FCV matrix. Each column in \mathbf{X} represents a FCV, and columns of \mathbf{W} are the vectorized representatives of AFCs. \mathbf{H} are the coefficients during matrix factorization. M denotes the dimension of FCV and AFC, and N denotes the number of subjects. The rank r denotes the number of bases in \mathbf{W}



Finally, the multiplicative updating rule is obtained as:

$$W_{i,j} \leftarrow W_{i,j} \frac{\sum_k X_{i,k} (\mathbf{W}^T \mathbf{X})_{j,k} / (\mathbf{W} \mathbf{W}^T \mathbf{X})_{i,k} + \sum_k X_{i,k} \sum_l W_{l,j} X_{l,k} / (\mathbf{W} \mathbf{W}^T \mathbf{X})_{l,k}}{\sum_k ((\mathbf{W}^T \mathbf{X})_{j,k} + \sum_l W_{l,j} X_{i,k})}. \quad (7)$$

It should be noted that \mathbf{W} must be initialized before we perform the multiplicative updating in Eq. (7). Random initialization is a commonly used strategy in unsupervised feature learning (Saxe et al. 2011; Kuncheva and Vetrov 2006; Cox and Pinto 2011). Importantly, random initialization can also do well as pre-trained initialization (Saxe et al. 2011). Thus, in this work, \mathbf{W} is randomly initialized for our AFCs learning.

A critical issue in any decomposition analysis is the selection of the number of factors (r). Brunet et al. (2004) proposed the cophenetic correlation coefficient (CCC) approach to determine the optimal value of r , and Hutchins et al. (2008) presented a variation of the residual sum of squares (RSS) to estimate the optimal value of r . The key concept in the CCC approach is the consensus matrix \mathbf{C} , which is the average of the clustering connectivity matrices over many runs (due to the random initializations, we cannot exactly obtain the same clustering connectivity matrix on each run). The clustering connectivity matrix is a binary matrix which indicates whether two samples belong to the same cluster. The entry $c_{i,j}$ in the clustering connectivity matrix is one if samples i and j belong to the same cluster, otherwise $c_{i,j}$ is zero (Lancichinetti and Fortunato 2012). Thus, the dispersion of the consensus matrix between 0 and 1 measures the robustness and reproducibility of the clustering results with respect to random initializations. The CCC is measured by the Pearson correlation of two distant matrices: the distance between samples induced by the consensus matrix, $\mathbf{I}-\mathbf{C}$, and the distance between samples induced by the linkage used in the reordering of \mathbf{C} (Brunet et al. 2004). The optimal value of r is where the magnitude of the CCC falls. It should be noted that if the magnitude of the

CCC falls at more than one rank, the optimal value of r is where the magnitude of the CCC drops off most steeply. RSS is the residual sum of square errors between the data matrix (\mathbf{X}) and the approximated matrix ($\mathbf{W} \mathbf{W}^T \mathbf{X}$). If r is equal to or surpasses its optimal value, the additional reduction of RSS is minor (Hutchins et al. 2008). During the determination of r , we also want to evaluate the robustness and reproducibility of the clustering results with respect to the random initialization. Thus, we adopted the CCC to estimate the optimal rank r . Here, we experimentally found that 50 runs were sufficient for the consensus matrix to converge in our applications and each run takes approximately 2 min on a computer with Intel Core i7@3.33GHz and 4GB DDR.

Results

In this section, we performed AFC learning on the individual FCV matrices which were constructed from the fMRI data of the MCI group and the NC group, respectively. Then, the approximation residual errors of each FCV from its approximations based on the learnt AFCs were computed for classification purpose. Apart from the optimal rank estimated by CCC, we also tried other ranks to study the effect of model selection during AFC learning. Finally, we compared our AFCs learnt by using the D-PNMF approaches with the ones learnt by using the K-SVD algorithm, which is another variation of matrix factorization but without non-negative constraints. Meanwhile, the comparison with the traditional clustering algorithm, k -means, was also provided.

AFC learning on the fMRI data of individual MCI/NC groups

In order to learn the AFCs in MCI and NC respectively, the datasets were divided into two groups, one MCI group including all the 22 MCI patients, and another NC group including all the 29 controls. Based on the methods described in section of “[functional connectome construction and vectorization](#)”, we can derive three FCV matrices of the MCI group ($\mathbf{X}_{63903 \times 22}^{(MCI)}$), NC group ($\mathbf{X}_{63903 \times 29}^{(NC)}$), and pooled MCI and NC groups ($\mathbf{X}_{63903 \times 51}^{(all)}$) (used for rank determination). First, the CCCs of the FCV matrix of pooled MCI and NC groups were computed for ranks r from 2 to 10, as shown in Fig. 5. As there is only two groups used in this work, i.e., MCI and NC, the testing range for r from 2 to 10 is sufficient in this work. The CCC falls most steeply at $r=2$. Thus, the optimal rank r is estimated as 2. Then, the D-PNMF approaches were applied to the two FCV matrices of the MCI group, $\mathbf{X}_{63903 \times 22}^{(MCI)}$, and NC group, $\mathbf{X}_{63903 \times 29}^{(NC)}$, with the estimated rank $r=2$ separately, and four AFCs derived, two from the MCI group, $\mathbf{W}_{63903 \times 2}^{(MCI)}$, and another two from the NC group, $\mathbf{W}_{63903 \times 2}^{(NC)}$.

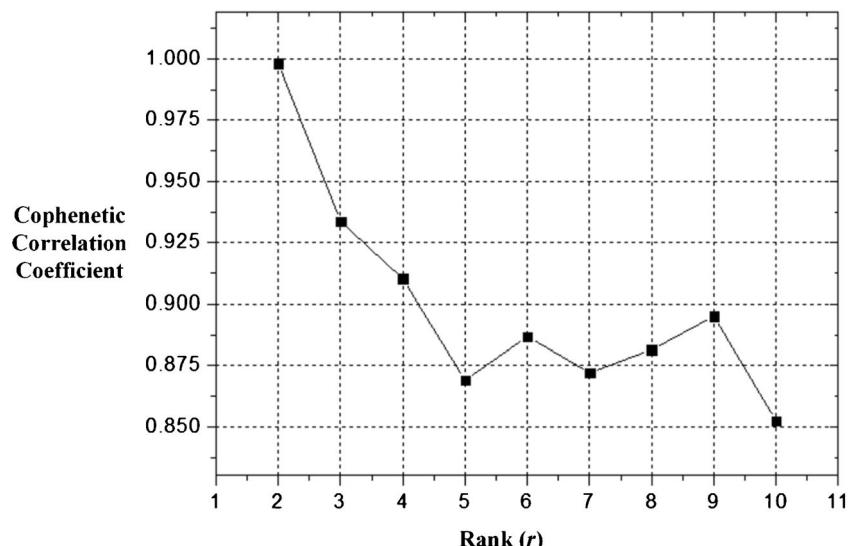
Then, the top 400 highest functional connectivity pairs in the four AFCs are experimentally selected to be visualized on the cortical surfaces, as shown in Fig. 6e–h. Notably, the first two AFCs in MCI and NC, i.e., MCI-AFC#1 (Fig. 6e) and NC-AFC#1 (Fig. 6g), are very similar to each other. This similarity is measured by the number of common connections (edges) on the cortical surfaces. There are 290 common connections between MCI-AFC#1 and NC-AFC#1, and the most connected ROIs are almost the same. While the last two AFCs in MCI and NC, i.e., MCI-AFC#2 (Fig. 6f) and NC-AFC#2 (Fig. 6h) exhibit significant differences from each other. Therefore, the four AFCs can be categorized into two pairs, one common pair composed of MCI-AFC#1 and NC-AFC#1, which indicates the common connectivity patterns within non-disease-specific sub-networks, and another distinctive pair

composed of MCI-AFC#2 and NC-AFC#2, which reflects the altered connectivity patterns within disease-specific sub-networks. The neuroscientific meanings of these four AFCs can be interpreted as follows. The common pair involves connectivities in the default mode network (DMN) (Yuan et al. 2013; Fox and Raichle 2007; Broyd et al. 2009), as highlighted by the yellow circles in Fig. 6e and g, and many strong inter-hemisphere connections in the occipital lobes as well. To better reveal the increased and decreased connectivity patterns within the distinctive pair, the functional connectivity strengths of MCI-AFC#2 and NC-AFC#2 were computed. The functional connectivity strength of each DICCOL was calculated by accumulatively summing all the functional connectivities between it and all of the other DICCOLs, as shown in Fig. 7b. Although both MCI-AFC#2 and NC-AFC#2 show strong connectivities in the prefrontal cortex, increased and decreased connectivity patterns in MCI-AFC#2, which possess increased and decreased function connectivity strengths, were clearly observed as compared with NC-AFC#2. MCI-AFC#2 exhibited hyper-connectivities in the right prefrontal cortex, left posterior parietal cortex and temporal lobes, while hypo-connectivities were exhibited in the dorsal parietal lobes and primary visual cortex, as highlighted by the yellow oval shapes in Fig. 7. The distinctive pair could be potentially considered as neuroimaging biomarkers for MCI in the future, and could be viewed as “atomic connectomics signatures”, which is one of the major contributions of this work.

Classification analysis

Based on the methods in the section of “[functional connectome construction and vectorization](#)” and results in the section of “[AFC learning on the fMRI data of individual MCI/NC groups](#)”, a FCV, \mathbf{V} , was constructed for each subject and two sets of bases

Fig. 5 The CCCs of the FCV matrix of the pooled MCI and NC group. The CCC falls most steeply at $r=2$. Thus, the rank r is estimated as 2



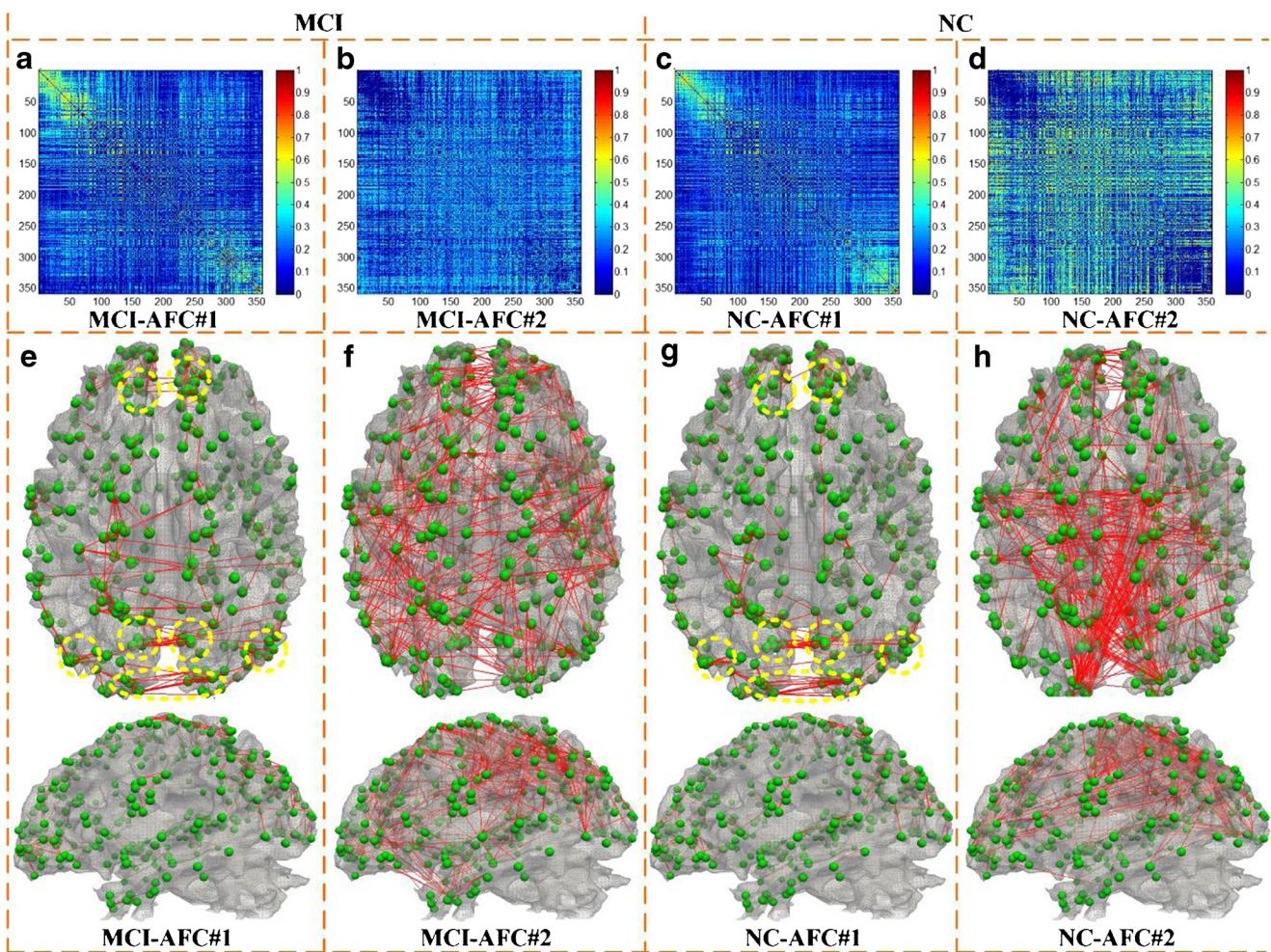


Fig. 6 Visualization of connectivity patterns of the four AFCs and of their projections on the WM/GM cortical surfaces in the top and lateral views. Due to the different coefficients during D-PNMF, the values of these matrices were very small. Thus, these matrices are scaled to the range of 0 to 1. Matrices are color-coded according to the strength of

functional connectivity. DICCCOL ROIs are marked as green spheres on the cortical surfaces, and the functional connectivities between ROIs are shown as red edges connecting those spheres. The top 400 highest functional connectivity pairs are selected for visualization. The yellow circles in (e) and (g) highlight the ROIs involved in the DMN

(i.e., AFCs) were learned from the MCI and NC groups, denoted as \mathbf{W}_{MCI} and \mathbf{W}_{NC} . From a mathematical perspective, the two sets of bases can span two spaces, each of which can be viewed as a class (MCI or NC). As discussed in section of “functional connectome construction and vectorization”, the projections of \mathbf{V} onto the two spaces, i.e., the approximations of \mathbf{V} based on the two sets of bases, could be computed as $\mathbf{V}' = \mathbf{W}^T \mathbf{W} \mathbf{V}$. Then, the approximation residual errors can be computed based on Eq. (2), as follows:

$$R = \|\mathbf{V} - \mathbf{W}^T \mathbf{W} \mathbf{V}\| \quad (8)$$

where $\|\cdot\|$ is a matrix norm. In this work, we used the ℓ^2 norm, and then R can be viewed as the sum of square errors between the FCV and its approximation. Given the two sets of bases, two approximation residual errors can be obtained for each subject, one based on \mathbf{W}_{MCI} (R_{MCI}) and another based on

\mathbf{W}_{NC} (R_{NC}), as shown in Fig. 8. The approximation residual error R evaluates how well the set of bases can characterize the FCV and could be used for classification purposes based on the least square error criterion (Mirkin 1998). If R_{MCI} is less than R_{NC} , meaning \mathbf{W}_{MCI} can better characterize the FCV of the test subject, the subject is classified as a MCI patient. On the contrary, if R_{NC} is less than R_{MCI} , the subject is classified as a NC subject. In this work, it was observed that R_{MCI} was not equal to R_{NC} for every subject.

In this work, there were only 22 MCI patients and 29 NC subjects. To examine the robustness and reproducibility of the classification results, the commonly used leave-one-out cross-validation strategy was adopted to evaluate the sensitivity and specificity of our learned AFCs. Finally, 100 % of MCI patients and 94.1 % of NC subjects in Dataset 1 were successfully classified, while 83.3 % of MCI patients 83.3 % were successfully classified in Dataset 2.

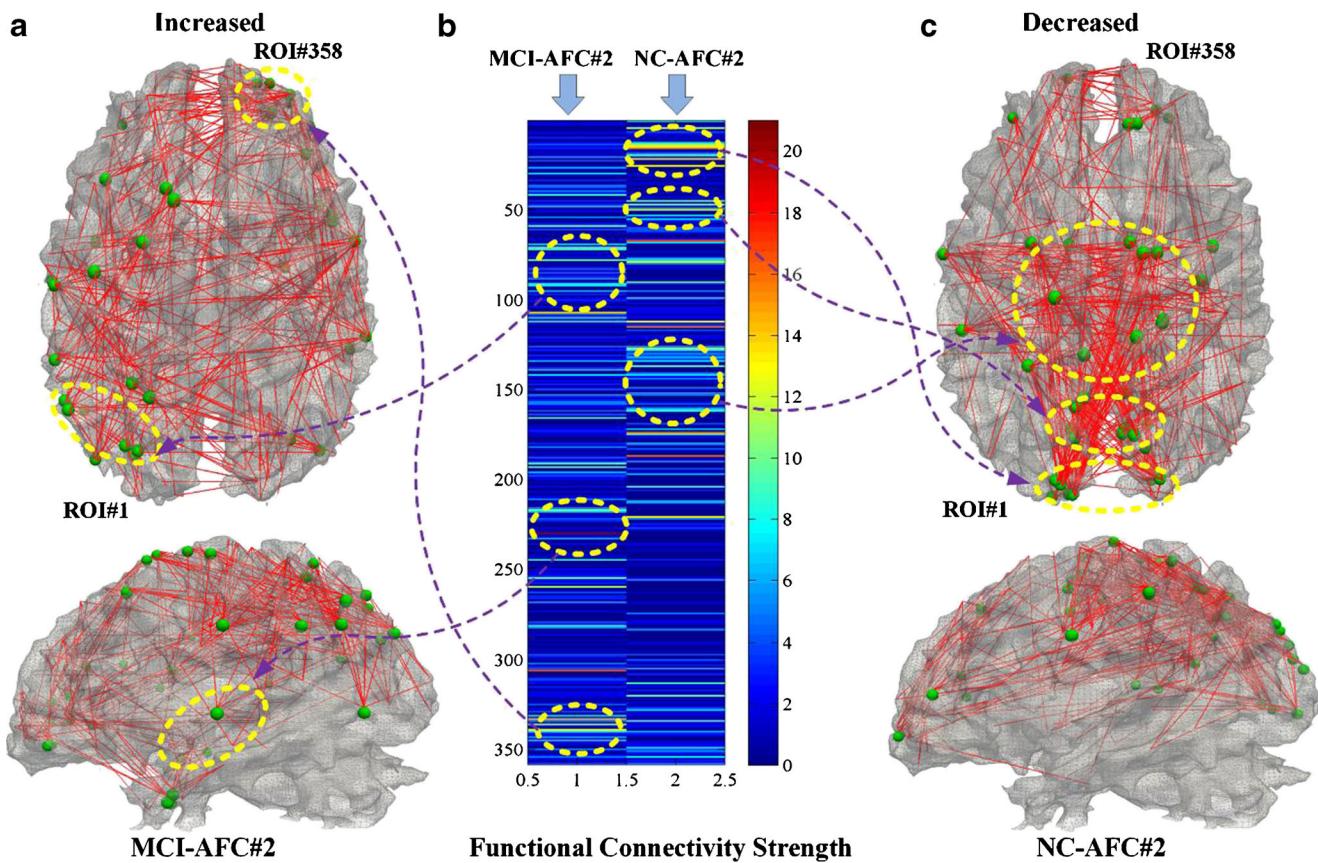


Fig. 7 Visualization of the functional connectivity strengths of MCI-AFC#2 and NC-AFC#2, as shown in (b). Those DICCCOLs with increased functional connectivity strength in MCI-AFC#2 are selected and visualized with the functional connectivity pattern of MCI-AFC#2, as shown in (a). Those DICCCOLs with decreased functional connectivity

strength in MCI-AFC#2 are selected and visualized with the functional connectivity pattern of NC-AFC#2 for better illustration, as shown in (c). In the DICCCOL system, the ROIs are indexed left-to-right and bottom-to-top in the top view

Studies on the effect of model order selection

To investigate the model sensitivity of the rank in D-PNMF, we tested the model by using various ranks on the same datasets

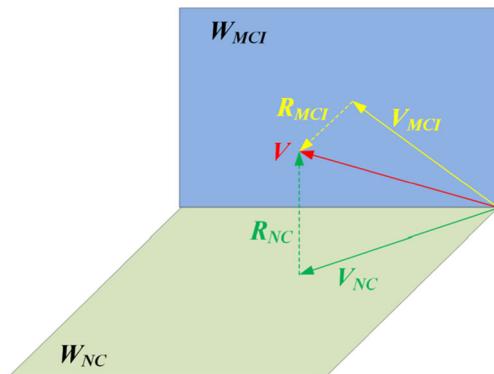


Fig. 8 The demonstration of the classification method based on the approximation residual error. The light blue area represents the MCI group characterized by \mathbf{W}_{MCI} , while the light olive area represents the NC group characterized by \mathbf{W}_{NC} . The yellow dotted lines indicate the approximation residual error R_{MCI} of V from its approximation V_{MCI} reconstructed from \mathbf{W}_{MCI} , while the green dotted lines indicate the approximation residual error R_{NC} of V from its approximation V_{NC} reconstructed from \mathbf{W}_{NC}

used in section of “AFC learning on the fMRI data of individual MCI/NC groups”. For each experiment, we performed AFC learning via the D-PNMF method with different ranks. We then compared the newly learned AFCs with the four AFCs described in Fig. 6, which are considered as the basis patterns in this experiment. By increasing the value of the rank r used in the D-PNMF method and conducting the comparison between the newly obtained AFCs and basis patterns, we observed that the additional AFCs were first split from the distinctive AFCs (i.e., MCI-AFC#2 and NC-AFC#2), and then from the common AFCs (i.e., MCI-AFC#1 and NC-AFC#1). More specifically, the common AFCs in NC (i.e., NC-AFC#1 in Fig. 6e) was split into two AFCs when the rank was increased to 4, while the common AFCs in MCI (i.e., MCI-AFC#1 in Fig. 6g) began to split when the rank was increased as large as 8. This suggested that the AFCs within non-disease-specific subnetworks in MCI are more stable than the ones in NC.

Comparisons with dictionary learning and k -means methods

With the objective function, $\|\mathbf{X} - \mathbf{WH}\|_p \leq \varepsilon$, where $\|\cdot\|_p$ is ℓ^p norm for $p=1, 2$ and ∞ , the approximation $\mathbf{X} \approx \mathbf{WH}$ can also be

used for dictionary learning (DL) (Aharon et al. 2006; Wright et al. 2009; Qiang and Baoxin 2010). Thus, we performed dictionary learning on the same datasets for comparison purposes. In this work, we applied an efficient implementation of the K-SVD algorithm based on batch orthogonal matching pursuit (OMP) (Aharon et al. 2006; Rubinstein et al. 2008; Pati et al. 1993) to these FCV matrices, with the dictionary size k of 2 for comparison purposes. Then, four sub-dictionaries were learned, two from the MCI group and another two from the NC group, as illustrated in Fig. 9. MCI-D#1 (Fig. 9e) and NC-D#1 (Fig. 9g) were somewhat similar to the common AFCs described in Fig. 6. It should be mentioned that there are negative entries in the connectivity matrix of the learned sub-dictionaries. The reason is that, without non-negative constraints, dictionary learning allows not only additive, but also subtractive, combinations. It means that we may learn non-zero entries in the sub-dictionaries, even if their corresponding entries in the FCV matrix are zeroes. If the corresponding entries in the FCV matrix are zeroes, it means

there are no co-variations between these corresponding ROIs, so that the corresponding non-zero entries in the sub-dictionaries are meaningless, which goes against our goals in this work.

Apart from dictionary learning, we also compared the D-PNMF approaches with traditional clustering algorithm. Take the wildly used k -means method as an example. For comparison purpose, the cluster number was set as 2 as well. After clustering with the k -means algorithm, those functional connectomes belonging to the same clusters were averaged. The averaged functional connectomes can be visualized in Fig. 10. Although the connectivity patterns (Fig. 10a–d) of these averaged functional connectomes were remarkably different from the ones of the common AFCs (Fig. 6a and c) described in Fig. 6, their projections on the cortical surfaces share many common connections with the common AFCs and could be approximately characterized by the common AFCs. Those four averaged functional connectomes are somehow similar to each other. More specifically, MCI-FC#1 (Fig. 10e)

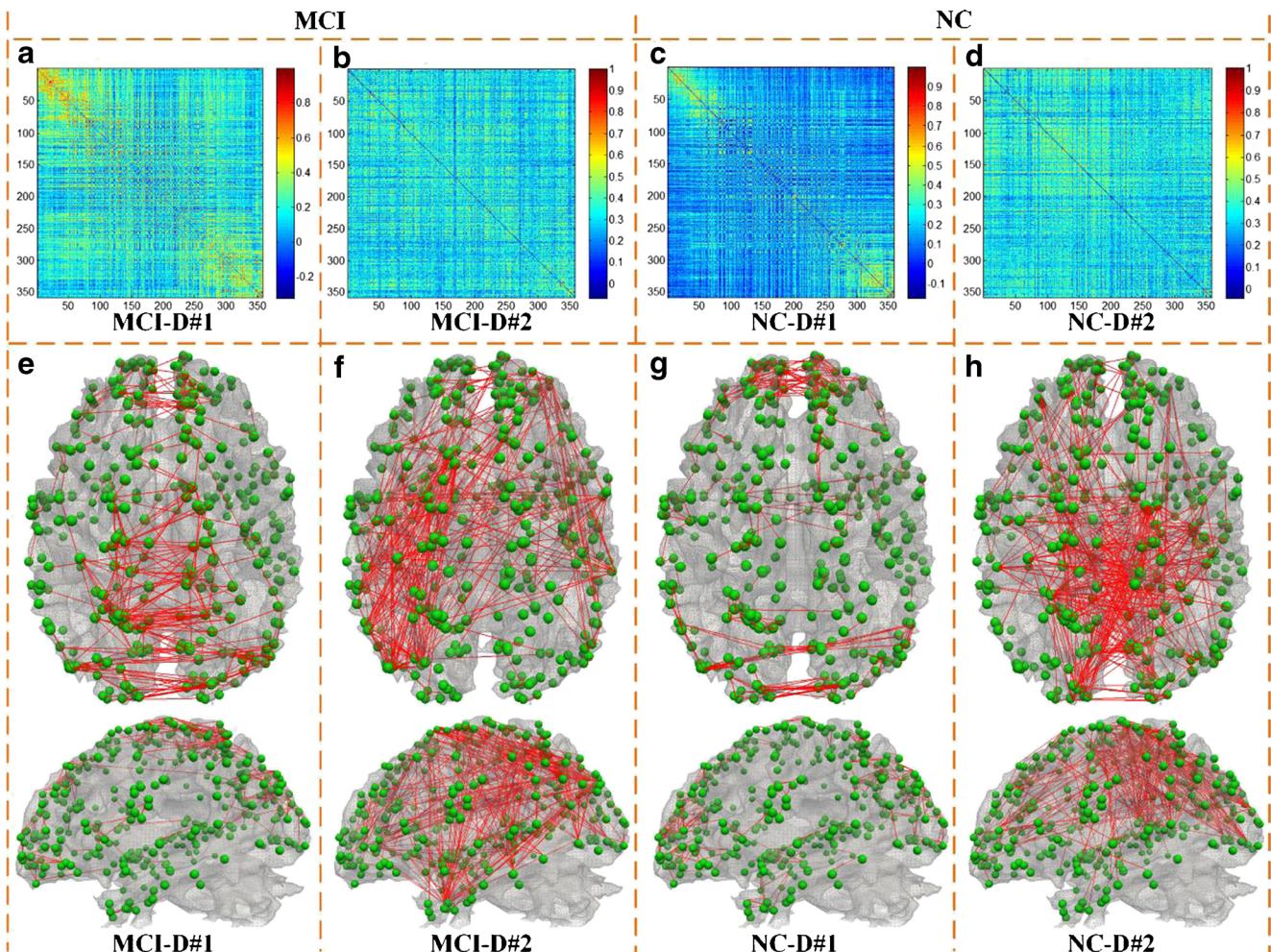


Fig. 9 Visualization of connectivity matrices of the four sub-dictionaries learned via the K-SVD algorithm based on batch OMP, and their projections on the cortical surfaces in the top and lateral views. Due to the

different coefficients of the four matrices during the K-SVD algorithm, the values of these matrices were also very small. Thus, these matrices are scaled to the range of 0 to 1

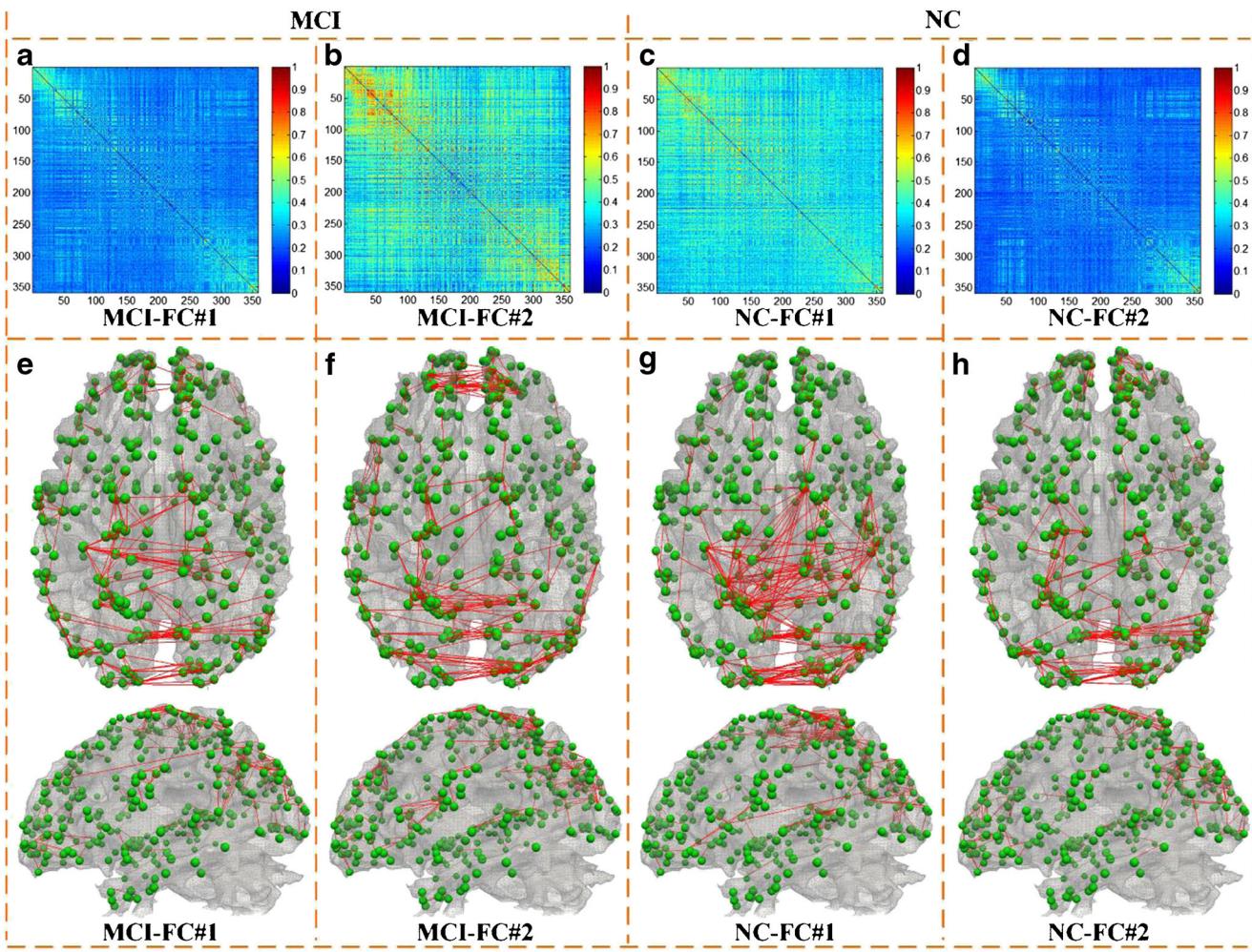


Fig. 10 Visualization of connectivity matrices derived by the k -means method (top panel) and of their projections on the WM/GM cortical surfaces (bottom panel). Although the functional connectivity patterns

in top panel are different, their projections on the cortical surfaces share many common connections

and MCI-FC#2 (Fig. 10f) share 329 and 168 common connections with MCI-AFC#1 (Fig. 6e) respectively, and NC-FC#1 (Fig. 10g) and NC-FC#2 (Fig. 10h) share 338 and 127 common connections with NC-AFC#1 (Fig. 6g) respectively. The experimental results suggest that our D-PNMF approaches can reveal more meaningful phenomena that can be revealed by traditional clustering methods.

Discussion and conclusion

Based on the consistent DTI-derived DICCCOL (Zhu et al. 2013) system, we constructed fine-granularity functional connectomes from the fMRI data of MCI and NC groups, and then achieved four AFCs by applying the D-PNMF approaches to the MCI and NC groups separately. By connectomics-level comparison, it was found that the four AFCs can be categorized into two pairs according to the

similarity among them, one common pair indicating the common connectivity patterns within non-disease-specific sub-networks in MCI and NC, and another distinctive pair implicating the altered connectivity patterns within disease-specific sub-networks in MCI, which can be viewed as “atomic connectomics signatures”. By further comparison based on the functional connectivity strengths of the distinctive pair, clearly increased and decreased connectivity patterns within disease-specific sub-networks were observed (Fig. 7). In addition, the four AFCs also had high classification accuracy including both sensitivity and specificity. Comparison studies show that the D-PNMF approaches can learn well part-based functional connectivity patterns in sub-networks, as compared with the K-SVD algorithm, which is another variant of matrix factorization but without non-negative constraints, and reveal more meaningful phenomena, as compared with the k -means algorithm.

Notably, we imposed the non-negative constraints on functional connectomes, i.e., we used the absolute values of the

Pearson correlations to evaluate the functional connectivities in connectomes. Actually, the positive Pearson correlations between fMRI BOLD signals measure the synchronous co-variations between ROIs, while the negative ones measure the asynchronous co-variations between ROIs. In the connectomics-level analysis of functional connectome on the cortical surfaces, we aim to analyze both synchronous and asynchronous co-variations between ROIs simultaneously. On the other hand, it is feasible to drop the negative sign according to information theoretical measurements. Another potential limitation is that the learned AFCs cannot be directly compared, due to the different coefficients associated with the AFCs during matrix factorization. In this study, the functional connectomes were compared at the connectomics level, i.e., the comparisons between the AFCs are based on their connections (edges) on the cortical surfaces. We experimentally chose the top 400 highest connectivity pairs in the AFCs to be visualized on the cortical surfaces for connectomics-level analysis and comparison.

Here, we employed the DICCCOLs as network nodes for functional connectome construction. It should be mentioned that the DICCCOL ROIs were defined as 3D grid vertices within the boundary box of the WM cortical surface, and was visualized on the WM/GM cortical surface. The WM/GM cortical surface is reconstructed based on the brain tissue map derived from DTI data using in-house software (Liu et al. 2004). Although it is more preferable that the DICCCOLs are placed on the “deep” gray matter, such as on the central cortical surface, given the 2 mm isotropic DTI data, it is hard to reconstruct the central cortical surface. Even though it is feasible to reconstruct the central cortical surface from structural MRI data, the mapping of such central cortical surface from MRI data to DTI data will be substantially distorted by the intrinsic geometric distortions in DTI data acquisition. Therefore, at this stage with current DTI data acquisition methods, we placed the landmarks on the inner GW/WM cortical surface, instead of the central cortical surface. Then a GM tissue map was used to mask those GM voxels such that only GM fMRI signals were extracted within the neighborhood of the DICCCOL landmarks.

Another potential limitation of this work is that the current 358 DICCCOLs are not the whole-cortex parcellation, but are sampled consistent landmarks (i.e. ROIs). Specifically, only a portion (2.5 %) of the total GM voxels are covered by the DICCCOL landmarks and used in the connectivity mapping. With our recently achieved DICCCOL-based whole-cortex parcellation, many more gray matter voxels could be included into the connectivity mapping, thus significantly alleviated the abovementioned problem. Our recent ongoing studies have also shown that the fMRI BOLD signals sampled by DICCCOL landmarks are quite representative of the whole-cortex signals in terms of the inference of concurrent brain network activities. Due to the scope of this paper, those results

will be reported in our future papers. In resting state, the spontaneous BOLD signals fluctuate in a low frequency, mainly between 0.01 Hz and 0.1 Hz (Fransson 2005; Fox and Raichle 2007). This is also the typical band that resting state networks are of interest. Thus, we performed band pass filtering (0.01 Hz~0.1 Hz) on the fMRI data.

In this work, the demographic information, such as age, education and cognitive performance, of the MCI group and the NC group were different in a single dataset and cross the two datasets. For example, MCI patients performed better than NC for MMSE in the dataset 1, while the MCI group is older than the NC group in the dataset 2. However, recall that our work is feasible to fuse and combine different datasets from different labs. These differences could somehow be neutralized by combining the two datasets. Moreover, we could introduce more datasets to deal with these potential issues in the future.

Notably, previous studies (Wee et al. 2012; Zhu et al. 2014) attempted to identify MCI patients from NC subjects and already achieved a relative high classification accuracy. Wee et al. (2012) identified about 70 % MCI subjects and 70.59 % NC subjects using connectivity networks constructed from the fMRI data in the dataset 1. Zhu et al. (2014) adopted a two-stage supervised feature selection procedure and achieved significantly high classification accuracy (more than 95 %). However, the fine-granularity connectivity networks have not been explored and visualized yet in those previous studies. Besides, many features were discarded during feature selection in (Zhu et al. 2014), which may be useful for analysis. Along with the high classification accuracy (more than 83 %), the fine-granularity connectivity networks with high differentiation power were also explored and visualized in this work. Importantly, all the features, not only the disease-related but also the non-disease-related, were preserved.

In this work, we experimentally adopted the Pearson correlation to estimate functional connectivities. We plan to investigate and compare other metrics such as the partial correlation, mutual information, and wavelet transform coherence (WTC) for the datasets in this paper in the future. We would like to point out that in our prior study (Zhang et al. 2013), Pearson correlation and partial correlation have comparable performances in classification applications. Furthermore, the work presented in this paper can also be further enhanced with a larger rank used in D-PNMF. The AFC learning was performed with a low rank of 2. And only two AFCs were obtained from each group. In general, it is still far from a full understanding of connectomics signatures in MCI in a finer granularity scale. In the future, we will focus on discovery of more connectomics signatures and assessment of meaningful connectivity patterns within more brain sub-networks. Also, more in-depth analyses of the neuroscientific meanings of these altered connectivity patterns should be performed by considering clinical and behavior datasets in the future. In

addition, the same computational framework can be possibly applied in many other psychiatric disorders to reveal the potentially dysfunctions of brain sub-networks.

Acknowledgments T Liu was supported by NIH R01 DA-033393, NIH R01 AG-042599, NSF CAREER Award IIS-1149260, NSF CBET-1302089 and NSF BCS-1439051. J Zhang was supported by start-up funding and Seseeal Award from Yale University. The authors would like to thank the anonymous reviewers for their constructive comments.

Conflict of Interest Jinli Ou, Li Xie, Xiang Li, Daqiang Zhu, Douglas P. Terry, A. Nicholas Puente, Rongxin Jiang, Yaowu Chen, Lihong Wang, Dinggang Shen, Jing Zhang, L. Stephen Miller, and Tianming Liu declare that they have no conflicts of interest.

Informed Consent All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, and the applicable revisions at the time of the investigation. Informed consent was obtained from all patients for being included in the study.

References

- Aharon, M., Elad, M., & Bruckstein, A. (2006). K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. *Signal Processing, IEEE Transactions on*, 54(11), 4311–4322.
- Arbabshirani, M. R., Kiehl, K. A., Pearson, G. D., & Calhoun, V. D. (2013). Classification of schizophrenia patients based on resting-state functional network connectivity. *Frontiers in Neuroscience*, 7, doi:10.3389/fnins.2013.00133.
- Bassett, D. S., & Bullmore, E. (2006). Small-world brain networks. *The Neuroscientist*, 12(6), 512–523.
- Broyd, S. J., Demanuele, C., Debener, S., Helps, S. K., James, C. J., & Sonuga-Barke, E. J. S. (2009). Default-mode brain dysfunction in mental disorders: a systematic review. *Neuroscience & Biobehavioral Reviews*, 33(3), 279–296.
- Brunet, J.-P., Tamayo, P., Golub, T. R., & Mesirov, J. P. (2004). Metagenes and molecular pattern discovery using matrix factorization. *Proceedings of the National Academy of Sciences*, 101(12), 4164–4169.
- Camchong, J., MacDonald, A. W., 3rd, Bell, C., Mueller, B. A., & Lim, K. O. (2011). Altered functional and anatomical connectivity in schizophrenia. *Schizophrenia Bulletin*, 37(3), 640–650.
- Cocchi, L., Bramati, I. E., Zalesky, A., Furukawa, E., Fontenelle, L. F., Moll, J., et al. (2012). Altered functional brain connectivity in a non-clinical sample of young adults with attention-deficit/hyperactivity disorder. *The Journal of Neuroscience*, 32(49), 17753–17761.
- Cox, D., & Pinto, N. (2011). Beyond simple features: A large-scale feature search approach to unconstrained face recognition. *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 I.E. International Conference on*, (pp. 8–15).
- Derrfuss, J., & Mar, R. A. (2009). Lost in localization: the need for a universal coordinate database. *NeuroImage*, 48(1), 1–7.
- Faraco, C. C., Puente, A. N., Brown, C., Terry, D. P., & Stephen Miller, L. (2013). Lateral temporal hyper-activation as a novel biomarker of mild cognitive impairment. *Neuropsychologia*, 51(11), 2281–2293.
- Fornito, A., & Bullmore, E. T. (2014). Connectomics: a new paradigm for understanding brain disease. *European Neuropsychopharmacology*, doi:10.1016/j.euroneuro.2014.02.011.
- Fox, M. D., & Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews. Neuroscience*, 8(9), 700–711.
- Fransson, P. (2005). Spontaneous low-frequency BOLD signal fluctuations: an fMRI investigation of the resting-state default mode of brain function hypothesis. *Human Brain Mapping*, 26(1), 15–29.
- Gilboa, A., Shalev, A. Y., Laor, L., Lester, H., Louzoun, Y., Chisin, R., et al. (2004). Functional connectivity of the prefrontal cortex and the amygdala in posttraumatic stress disorder. *Biological Psychiatry*, 55(3), 263–272.
- Greicius, M. D., Srivastava, G., Reiss, A. L., & Menon, V. (2004). Default-mode network activity distinguishes Alzheimer's disease from healthy aging: Evidence from functional MRI. *Proceedings of the National Academy of Sciences of the United States of America*, 101(13), 4637–4642.
- Hughes, C. P., Berg, L., Danziger, W. L., Coben, L. A., & Martin, R. L. (1982). A new clinical scale for the staging of dementia. *The British Journal of Psychiatry*, 140(6), 566–572.
- Hutchins, L. N., Murphy, S. M., Singh, P., & Gruber, J. H. (2008). Position-dependent motif characterization using non-negative matrix factorization. *Bioinformatics*, 24(23), 2684–2690.
- Kennedy, D. (2010). Making connections in the connectome era. *Neuroinformatics*, 8(2), 61–62.
- Kuncheva, L. I., & Vetrov, D. P. (2006). Evaluation of stability of k-Means cluster ensembles with respect to random initialization. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(11), 1798–1808.
- Lancichinetti, A., & Fortunato, S. (2012). Consensus clustering in complex networks. *Scientific Reports*, 2. doi:10.1038/srep00336.
- Lanius, R. A., Williamson, P. C., Bluhm, R. L., Densmore, M., Bokslman, K., Neufeld, R. W. J., et al. (2005). Functional connectivity of dissociative responses in posttraumatic stress disorder: a functional magnetic resonance imaging investigation. *Biological Psychiatry*, 57(8), 873–884.
- Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755), 788–791.
- Li, K., Guo, L., Zhu, D., Hu, X., Han, J., & Liu, T. (2012). Individual functional ROI optimization via maximization of group-wise consistency of structural and functional profiles. *Neuroinformatics*, 10(3), 225–242.
- Li, K., Zhu, D., Guo, L., Li, Z., Lynch, M. E., Coles, C., et al. (2013). Connectomics signatures of prenatal cocaine exposure affected adolescent brains. *Human Brain Mapping*, 34(10), 2494–2510.
- Li, X., Zhu, D., Jiang, X., Jin, C., Zhang, X., Guo, L., et al. (2014). Dynamic functional connectomics signatures for characterization and differentiation of PTSD patients. *Human Brain Mapping*, 35(4), 1761–1778.
- Liu, T. (2011). A few thoughts on brain ROIs. *Brain Imaging and Behavior*, 5(3), 189–202.
- Liu, T., Shen, D., & Davatzikos, C. (2004). Deformable registration of cortical structures via hybrid volumetric and surface warping. *NeuroImage*, 22(4), 1790–1801.
- Mirkin, B. (1998). *Mathematical classification and clustering: From how to what and why. Classification, data analysis, and data highways* (pp. 172–181). Berlin: Springer.
- Ou, J., Lian, Z., Xie, L., Li, X., Wang, P., Hao, Y., et al. (2014). Atomic dynamic functional interaction patterns for characterization of ADHD. *Human Brain Mapping*, 35(10), 5262–5278.
- Passingham, R. E., Stephan, K. E., & Kotter, R. (2002). The anatomical basis of functional localization in the cortex. *Nature Reviews. Neuroscience*, 3(8), 606–616.
- Pati, Y. C., Rezaifar, R., & Krishnaprasad, P. S. (1993). Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. *Signals, Systems, and Computers, Proceedings of the 27 th Annual Asilomar Conference on*, (pp. 40–44).
- Poldrack, R. A. (2012). The future of fMRI in cognitive neuroscience. *NeuroImage*, 62(2), 1216–1220.

- Puente, A. N., Faraco, C., Terry, D. P., Brown, C., & Miller, L. S. (2014). Minimal functional brain differences between older adults with and without mild cognitive impairment during the stroop. *Aging, Neuropsychology, and Cognition, 21*(3), 346–369.
- Qiang, Z., & Baoxin, L. (2010). Discriminative K-SVD for dictionary learning in face recognition. *Computer Vision and Pattern Recognition (CVPR), 2010 I.E. Conference on*, (pp. 2691–2698).
- Rubinstein, R., Zibulevsky, M., & Elad, M. (2008). Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit. *CS Technion*. doi:[10.1.1.182.9978](https://doi.org/10.1.1.182.9978).
- Santhanam, P., Coles, C. D., Li, Z., Li, L., Lynch, M. E., & Hu, X. (2011). Default mode network dysfunction in adults with prenatal alcohol exposure. *Psychiatry Research: Neuroimaging, 194*(3), 354–362.
- Saxe, A., Koh, P. W., Chen, Z., Bhand, M., Suresh, B., & Ng, A. Y. (2011). On random weights and unsupervised feature learning. *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, (pp. 1089–1096).
- Sporns, O. (2011). The human connectome: a complex network. *Annals of the New York Academy of Sciences, 1224*(1), 109–125.
- Sporns, O. (2013). Network attributes for segregation and integration in the human brain. *Current Opinion in Neurobiology, 23*(2), 162–171.
- Sporns, O., Chialvo, D. R., Kaiser, M., & Hilgetag, C. C. (2004). Organization, development and function of complex brain networks. *Trends in Cognitive Sciences, 8*(9), 418–425.
- Stam, C. J. (2010). Characterization of anatomical and functional connectivity in the brain: a complex networks perspective. *International Journal of Psychophysiology, 77*(3), 186–194.
- Supekar, K., Menon, V., Rubin, D., Musen, M., & Greicius, M. D. (2008). Network analysis of intrinsic functional brain connectivity in Alzheimer's disease. *PLoS Computational Biology, 4*(6), e1000100.
- Trefethen, L. N., & Bau III, D. (1997). *Numerical linear algebra* (Vol. 50): Siam.
- Venkataraman, A., Whitford, T. J., Westin, C. F., Golland, P., & Kubicki, M. (2012). Whole brain resting state functional connectivity abnormalities in schizophrenia. *Schizophrenia Research, 139*(1–3), 7–12.
- Wang, K., Liang, M., Wang, L., Tian, L., Zhang, X., Li, K., et al. (2007). Altered functional connectivity in early Alzheimer's disease: a resting-state fMRI study. *Human Brain Mapping, 28*(10), 967–978.
- Wee, C.-Y., Yap, P.-T., Zhang, D., Denny, K., Browndyke, J. N., Potter, G. G., et al. (2012). Identification of MCI individuals using structural and functional connectivity networks. *NeuroImage, 59*(3), 2045–2056.
- Wright, J., Yang, A. Y., Ganesh, A., Sastry, S. S., & Yi, M. (2009). Robust face recognition via sparse representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 31*(2), 210–227.
- Yang, Z., & Oja, E. (2010). Linear and nonlinear projective nonnegative matrix factorization. *Neural Networks, IEEE Transactions on, 21*(5), 734–749.
- Yang, Z., Yuan, Z., & Laaksonen, J. (2007). Projective non-negative matrix factorization with applications to facial image processing. *International Journal of Pattern Recognition and Artificial Intelligence, 21*(08), 1353–1362.
- Yuan, Y., Jiang, X., Zhu, D., Chen, H., Li, K., Lv, P., et al. (2013). Meta-analysis of functional roles of DICCCOLs. *Neuroinformatics, 11*(1), 47–63.
- Zhang, Y., Han, J., Hu, X., Guo, L., & Liu, T. (2013). Data-driven evaluation of functional connectivity metrics. *Biomedical Imaging (ISBI), 2013 I.E. 10th International Symposium on*, (pp. 532–535).
- Zhu, D., Li, K., Faraco, C. C., Deng, F., Zhang, D., Guo, L., et al. (2012). Optimization of functional brain ROIs via maximization of consistency of structural connectivity profiles. *NeuroImage, 59*(2), 1382–1393.
- Zhu, D., Li, K., Guo, L., Jiang, X., Zhang, T., Zhang, D., et al. (2013). DICCCOL: dense individualized and common connectivity-based cortical landmarks. *Cerebral Cortex, 23*(4), 786–800.
- Zhu, D., Li, K., Terry, D. P., Puente, A. N., Wang, L., Shen, D., et al. (2014). Connectome-scale assessments of structural and functional connectivity in MCI. *Human Brain Mapping, 35*(7), 2911–2923.