Sci. Bull. (2015) 60(1):86–95
DOI 10.1007/s11434-014-0698-3

www.scibull.com
www.springer.com/scp

Article

Life & Medical Sciences

A Connectome Computation System for discovery science of brain

Ting Xu · Zhi Yang · Lili Jiang · Xiu-Xia Xing · Xi-Nian Zuo

Received: 24 November 2014 / Accepted: 1 December 2014 / Published online: 5 January 2015

Abstract Much like genomics, brain connectomics has rapidly become a core component of most national brain projects around the world. Beyond the ambitious aims of these projects, a fundamental challenge is the need for an efficient, robust, reliable and easy-to-use pipeline to mine such large neuroscience datasets. Here, we introduce a computational pipeline—namely the Connectome Computation System (CCS)—for discovery science of human brain connectomes at the macroscale with multimodal magnetic resonance imaging technologies. The CCS is designed with a three-level hierarchical structure that includes data cleaning and preprocessing, individual connectome mapping and connectome mining, and knowledge discovery. Several functional modules are embedded into this hierarchy to implement quality control procedures, reliability analysis and connectome visualization. We demonstrate the utility of the CCS based upon a publicly available dataset, the NKI–Rockland Sample, to delineate the normative trajectories of well-known large-scale neural networks across the natural life span (6–85 years of age). The CCS has been made freely available to the public via GitHub (https://github.com/zuoxinian/CCS) and our laboratory's Web site (http://lfcd.psych.ac.cn/ccs.html) to facilitate progress in discovery science in the field of human brain connectomics.

Keywords Connectome · Life span · Big data · Normative charts · Discovery science

Ting Xu and Zhi Yang contributed equally to this work.

T. Xu · Z. Yang · L. Jiang · X.-N. Zuo (✉)
Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China
e-mail: zuoxn@psych.ac.cn; zuoxinian@gmail.com
URL: http://lfcd.psych.ac.cn

T. Xu · Z. Yang · L. Jiang · X.-N. Zuo
Magnetic Resonance Imaging Research Center, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China

T. Xu · Z. Yang · L. Jiang · X.-N. Zuo
Laboratory for Functional Connectome and Development, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China

X.-X. Xing
Department of Applied and Computational Mathematics, College of Applied Sciences, Beijing University of Technology, Beijing 100124, China

X.-N. Zuo
Faculty of Psychology, Southwest University, Chongqing 400715, China

1 Introduction

In connectomics, the human brain is graphically modeled as a complex system of nodes and edges between these nodes known as the connectome [1, 2]. This perspective has revolutionized brain research by combining complex network science methodology with advances in neuroimaging [3, 4]. It also calls classic computational modeling methods back to the forefront of the field (e.g., generative connectome models) [5–9]. These models remind brain scientists that the most amazing feature of the human connectome is how it functions (i.e., functional connectomics) [10–13]. Connectome projects have been initiated in many countries, energizing the field of neuroscience with the possibility that the mysteries of the brain could be unlocked by studying connectomic associations with individual differences in behavior and state of mind as well as in gene expression and cell signaling.

The growth of the fields of genomics and connectomics has been similar in magnitude, but with an inverse focus on big data interests. Genomics undertook its landmark project, the Human Genome Project (HGP), to determine the sequence of the chemical base pairs in human DNA and to identify and map all the genes at the level of the individual [14, 15]. More recently, big genome data have been obtained at the group level by efforts such as the 1,000 Genomes Project to provide a deep characterization of human genome sequence variation as a foundation for investigating the relationship between genotype and phenotype [16, 17]. In contrast, connectomics began with both grassroots and government projects at the group level—the 1,000 Functional Connectomes Project (FCP [10]) and the Human Connectome Project (HCP [18]), respectively—to investigate individual differences in large-scale brain networks and link these differences to human mental state and behavior. In 2013, the US government initiated the Brain Research through Advancing Innovative Neurotechnologies (BRAIN) project to produce an individual brain map showing how individual cells and complex neural circuits interact over both time and space [12].

Bioinformatics approaches provide the highly efficient, robust and reliable computational support necessary for the data mining and knowledge discovery of big genomics data [19, 20]. In a similar vein, the recent explosion of big neuroscience data [21] has pushed forward the development of similar resources to enable the discovery science of human brain function [10]. Neuroinformatics has thus emerged as a field to provide tools for data storage, management and sharing (SMS) and data mining, analysis and visualization (MAV) in human brain connectomics. SMS tools such as XNAT [22] and COINS [23] have greatly advanced this field by relatively standard methods of managing and sharing data. There are many MAV tools currently available. However, widely accepted standards are missing for developing MAV tools and related algorithms. Tools such as SPM, FSL, AFNI and FreeSurfer are designed to address multimodal brain imaging data and have overlapping functions of data analysis, which can serve as the basis for building robust and reliable pipelines for big connectomics data. Such a pipeline should be configurable, reliable and extendable for data mining and knowledge discovery in big connectomics data. Currently, most pipelines, such as REST, BCT, the HCP pipeline, PANDA and the Connectome Mapper, seem not to explicitly meet all three requirements (see Table 1 for a full list of software currently available with detailed information).

Inspired by the fact that the human brain connectome is organized in a highly hierarchical and modular network, we designed a Connectome Computation System (CCS) to fulfill these three (configurable, reliable and extendable) requirements. Specifically, a highly hierarchical and modular structure was developed in the CCS to make highly efficient computation feasible. Configuration files were used for all levels of computation across different hierarchies and modules. Modules with different functions were implemented across different levels of hierarchies to make novel algorithms of data mining and knowledge discovery easy to integrate. Here, we introduce an alpha version of CCS and demonstrate its advantages by charting the life span trajectories of seven common large-scale neural networks. We also discuss upcoming functions of the CCS intended to expedite human brain discovery science. We hope that the release of the CCS will contribute to and expedite open science of the human brain.

2 Overall design strategy

The CCS pipeline aims to be a feasible and reliable toolbox for data mining and knowledge discovery for macroscale connectomics [24]. An overall design strategy is proposed to meet this demand. It contains a three-level hierarchical structure including low-level data cleaning and preprocessing (H1), mid-level individual connectome mapping (H2) and high-level connectome mining and knowledge discovery (H3). Both intra- and inter-hierarchical modular designs are implemented for different functions. The CCS integrates currently available software and packages as much as possible to process analytic requests at any hierarchy. Figure 1 contains a flowchart for further delineating this overall design strategy.

CCS pipelines are mostly written in bash shell and MATLAB scripts by integrating many functions, which come from two types of resources: (1) publicly available neuroimaging toolboxes such as SPM, FSL, AFNI and FreeSurfer, and (2) in-house developed connectome algorithms. In the following sections, we detail these functions across different CCS hierarchies, with special emphasis on the connectome algorithms developed in our laboratory. All procedures followed were in accordance with the ethical standards of the committee on responsible human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2008. Informed consent was obtained from all participants before inclusion in the study.

3 H1: Data cleaning and preprocessing

This hierarchy represents the most common processing steps for cleaning and preprocessing multimodal MRI datasets and currently contains three modules designed for structural MRI (sMRI), diffusion MRI (dMRI) and resting-state functional MRI (rfMRI). These preprocessing steps are performed in both volume and surface space. Details of each module are described in the following subsections.

Table 1 Representative softwares for human brain connectomics

Software	Full name	Programming language	Data modality	Availability
AFNI	Analysis of Functional NeuroImages	C	tfMRI/rfMRI	http://afni.nimh.nih.gov/afni
BCT	Brain Connectivity Toolbox	MATLAB	Graphs	http://www.brain-connectivity-toolbox.net
BNV	Brain Net Viewer	MATLAB	Graphs	http://www.nitrc.org/projects/bnv
CCS	Connectome Computation System	Shell/MATLAB/R/Python	sMRI/dMRI/fMRI	http://lfcd.psych.ac.cn/ccs.html
CMTK	Connectome Mapping Toolkit	Python	sMRI/dMRI	http://www.cmtk.org
CPAC	Configurable Pipeline for the Analysis of Connectomes	Python	rfMRI	http://fcp-indi.github.io
CONN	Functional Connectivity Toolbox	MATLAB	fMRI	http://www.nitrc.org/projects/conn
DPARSF	Data Processing Assistant for Resting-State fMRI	MATLAB	rfMRI	http://rfmri.org/dparsf
FreeSurfer	FreeSurfer	C++/C/Shell	MRI	http://surfer.nmr.mgh.harvard.edu/fswiki
FSL	FMRIB Software Library	C++/Shell	MRI/ASL	http://fsl.fmrib.ox.ac.uk/fsl/fslwiki
GIFT	Group ICA of FMRI Toolbox	MATLAB	fMRI	http://mialab.mrn.org/software/gift
GraphVar	Graph Analysis of Brain Connectivity	MATLAB	fMRI	http://www.nitrc.org/projects/graphvar
GRETNA	Graph Theory Toolkit for Network Analysis	MATLAB	Graphs	http://www.nitrc.org/projects/gretna
HCP	Human Connectome Pipeline	Shell/MATLAB	MRI/MEG	https://github.com/Washington-University/Pipelines
NIAK	Neuroimaging Analysis Kit	MATLAB/OCTAVE	fMRI	http://www.nitrc.org/projects/niak
PANDA	Pipeline for Analyzing braiN Diffusion imAges	MATLAB	dMRI	http://www.nitrc.org/projects/panda
REST	Resting-State fMRI Analysis Toolkit	MATLAB	rfMRI	http://restfmri.net
SPM	Statistical Parametric Mapping	MATLAB/C	MRI/PET/EEG/MEG	http://www.fil.ion.ucl.ac.uk/spm

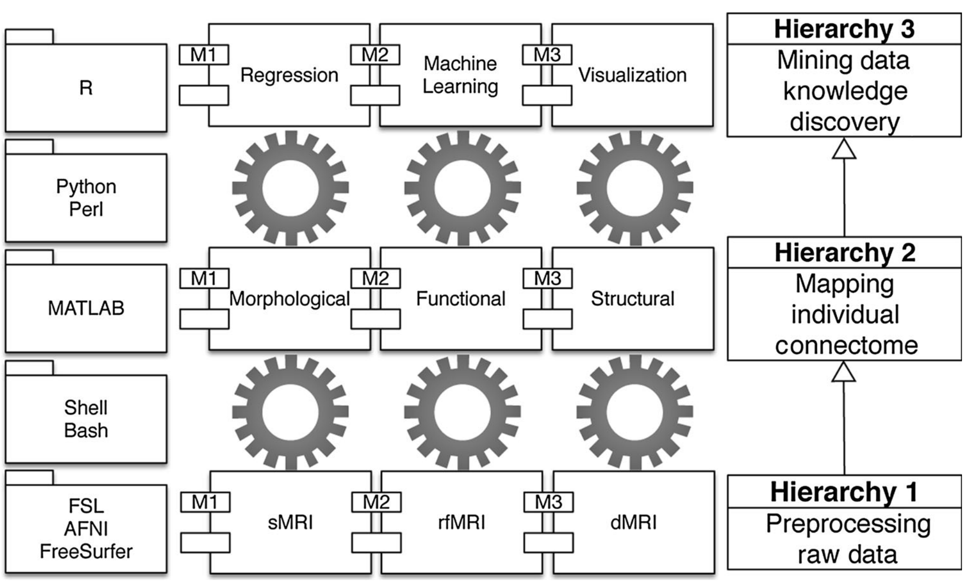

Fig. 1 Diagrammatic sketch of the overall design strategy. Three hierarchies support the CCS (right column). Each hierarchy contains several functional modules (e.g., M1, M2 and M3), which are implemented with various neuroimaging toolboxes and other scripts (left column)

3.1 Preprocessing sMRI data

sMRI processing aims to reconstruct the brain's cortical surfaces [25–28] and includes a series of steps: (1) spatial noise removal by a non-local means filtering operation [29, 30]; (2) image intensity inhomogeneity correction; (3) brain extraction using an improved hybrid watershed/surface deformation procedure; (4) tissue segmentation of the cerebrospinal fluid

(CSF), white matter (WM) and deep gray matter (GM); (5) disconnection of the two hemispheres and subcortical structures; (6) fixation of the interior holes of the segmentation; (7) a triangular mesh tessellation over the GM–WM boundary; (8) a mesh deformation to produce a smooth representation of the GM–WM interface (white surface) and GM–CSF interface (pial surface); (9) correction for topological defects on the surface; (10) individual surface mesh inflation into a sphere; and (11) individual surface (volume) normalization by estimating the deformation between the resulting spherical mesh (the individual brain volume) and a common spherical (volumetric) coordinate system that aligned the cortical folding patterns (image density similarities) across subjects.

Brain extraction (skull stripping) and registration (spatial normalization) are two challenging steps involved in preprocessing sMRI images. Our design allows the CCS to address these challenges by integrating the most recent advances in these fields. Currently, brain extraction is implemented by combining FSL and FreeSurfer with an optional step of manual intervention. In the future, this will be upgraded with knowledge-driven algorithms such as BEaST [31]. For registration, spatial normalization of volumetric images is performed with FLIRT/FNIRT in FSL and will soon be updated with the ANTs implementations [32].

3.2 Preprocessing dMRI data

The CCS employs FSL's Diffusion Toolbox (FDT [33]) to preprocess dMRI images. This includes: (1) correction for eddy current distortions and motion by realigning all images to the unweighted B0 images (or to the mean image if there are multiple B0 images) [34]; (2) tensor reconstruction by fitting the diffusion profile within each voxel; (3) estimation of the diffusion direction within each voxel as the principal eigenvector of the tensor by computing its eigen system [35]; (4) computation of fractional anisotropy (FA) in each voxel as the square root of the sum of squares (SRSS) of the diffusivity differences, divided by the SRSS of the diffusivities [36]; (5) reconstruction of the whole-brain white matter tracts with fiber assignment by continuous tracking (FACT) algorithm [37]; and (6) co-registration between dMRI and sMRI data by aligning the B0 brain image to the sMRI brain image by boundary-based registration (BBR) [38].

The fast and accurate reconstruction of the whole-brain white matter tracts is still a topic of debate. At present, the CCS chooses the fastest algorithm for this purpose. More comprehensive algorithms, such as probabilistic fiber tracking in FSL [39] and FATCAT in AFNI [40], can be easily integrated into the pipeline.

3.3 Preprocessing rfMRI data

The CCS functional preprocessing pipeline consists of: (1) discarding the first several volumes (total scanning duration of 10 s); (2) removing and interpolating temporal spikes from hardware instability or from head motion [41–43]; (3) correcting acquisition timing among image slices and head motion among image volumes; (4) normalizing the 4D global mean of image intensity to 10,000; (5) matching spatial correspondences between individual functional images and anatomical images by employing white surface boundary-based registration (BBR) algorithm [38]; (6) eliminating the effect of head motion and physiological noises during the rfMRI scanning session by regressing out the estimated Friston's 24-parameter motion curves [44, 45] and nuisance signals measured as WM and CSF mean time series from individual rfMRI time series [46]; (7) removing both linear and quadratic trends from the rfMRI data by multiple linear regression; and (8) projecting individual rfMRI time series onto the *fsaverage* surface grid and down-sampling to the *fsaverge5* surface grid (average inter-vertex distance = 4 mm) [47] or to the MNI152 standard volumetric space (spatial resolution = 3 mm).

The steps described above are common preprocessing steps in most rfMRI data analyses. However, some types of rfMRI analysis require different or additional preprocessing steps. For example, temporal filtering is usually performed for functional connectivity analyses but is not necessary for analyses in the frequency domain. Another example is global signal regression, which is a controversial preprocessing of rfMRI data [41, 48–50]. The CCS leaves the user to decide whether to use these two optional steps.

Following the processing, CCS provides a module to aid quality control. This quality control procedure (QCP) produces a set of screenshots for the user's visual inspection of the data and processing quality as well as a set of quality metrics for subsequent statistical analysis. Specifically, these screenshots are generated for the visual inspection of: (1) brain extraction or skull stripping; (2) brain tissue segmentation; (3) pial and white surface reconstruction; (4) BBR-based functional and diffusion image registration; and (5) head motion during rfMRI/dMRI. The quality metrics are computed to quantify: (1) the maximum distance of translational head movement (maxTrans); (2) the maximum degree of rotational head movement (maxRot); (3) the mean frame-wise displacement (meanFD); and (4) the minimal cost of the BBR co-registration (mcBBR). More quality metrics, which have been demonstrated in our recent paper about the Consortium for Reliability and Reproducibility (CoRR [51]), will be included in a future release of the CCS.

4 H2: Individual connectome mapping

Individual connectomes are constructed either explicitly (i.e., the structural connectome) or implicitly (i.e., the functional connectome) from their graph or connection matrices. Connectomic properties are characterized by a set of network metrics across the three (sMRI, dMRI and rfMRI) imaging modalities. Morphological connectomes are characterized by various morphological metrics (e.g., thickness, area, volume, curvature and gyrification index) of either large-scale units of brain parcellation or of vertices and voxels. These measurements can be further employed to construct a morphological brain graph [52–55] and quantitate the brain network with graphical metrics (e.g., centrality, efficiency, cost, module, rich club and small world) [56]. Similarly, structural connectomes are built of dMRI derivatives such as fibers, FA, MD, RD, AD, PD and L1 [57–59]. Limited by available imaging technology, dMRI-based connectomes can only be studied via large-scale brain parcellation.

Functional connectomics is a major feature of the current CCS pipeline. Two types of rfMRI-derived connectomics metrics are estimated in the CCS regarding whether the functional connectomes are constructed explicitly. Although it is currently the most widely used method, it usually is not accurate to generate a graph by defining its edges by internode statistical dependence (i.e., searching for significance) between time series. In particular, it is unclear whether too much or too little information has been omitted or how ignoring statistically insignificant edges by thresholding dependence (e.g., correlation) impacts the results. This leads to difficulties in interpreting the additional graphical metrics based on such a 'virtual' graph. In addition, we have recently highlighted the functional connectomics view of the human brain [60]. The brain is a connectome functioning across multiple temporal and spatial scales regardless of our interpretation of the data from an organism or of how we reconstruct its graphical representation. The best approach might be to apply or develop a network metric to measure the rfMRI time series directly, which is the functional output of the human connectome. Such a method would preserve all the information observed.

The CCS maps a series of functional features of the human brain connectome at levels of both vertex or voxel and large-scale parcellation. One advantage of these mapping processes is the implementation of surface-based computation [61, 62] by combining data from both the sMRI and rfMRI modules of the hierarchy 1 (H1). Here, we detail various functional connectomics maps with surface-based approaches, whereas volume-based maps can be estimated similarly for these functional metrics. Individual surface maps without an explicit reconstruction of the brain graph with rfMRI include: (1) amplitude of low-frequency fluctuations (ALFF) and its standardized version (fALFF) [63–65]; (2) local functional homogeneity or regional homogeneity (2dReHo [61, 66]); (3) voxel-mirrored homotopic connectivity (VMHC [67]); (4) seed region of interests functional connectivity (seedFC [68, 69]); and (5) independent component analysis with dual (spatial and temporal) regression (drICA) [70]. Individual surface maps with an explicit reconstruction of the brain graph by rfMRI include various vertex-wise network centrality maps (VNCM) [71]: (1) degree centrality (DC); (2) subgraph centrality (SC); (3) eigenvector centrality (EC); and (4) page-rank centrality (PC). All these vertex-wise maps are finally summarized into their large-scale parcellation distribution with both mean and standard deviation of each parcellation unit (parcel). The CCS employs two large-scale parcellation templates in FreeSurfer, Destrieux Anatomical Atlas [72] and Yeo2011_N1000 Functional Atlas [47] to summarize all these functional connectomics metrics at the parcel level.

The high efficiency of the CCS at mapping individual functional connectomics is mainly the result of two core CCS functions inspired by sparse and block matrix theory. *ccs_core_fastCoRR* is designed to quickly compute a Pearson's correlation coefficient and is based on a simple mathematical fact [71]: Given two standardized variables (mean $= 0$ and standard deviation $= 1$), the correlation between them is equivalent to their inner product. Thus, the sample correlation matrix between two large sample matrices X and Y is $\mathrm{R} = \mathrm{X}'\mathrm{Y}/(n-1)$, where n is the number of samples. This function is highly efficient at dealing with big neuroscience data in a high-performance computational environment (e.g., a large amount of physical memory). Neuroscientists who face the challenges of the big data of functional interactions across multiple scales of time and space have increasingly recognized the importance [21, 73] of a relaxed version of this method for use on personal computers [74]. To meet this need, the CCS turns sparse block matrix theory into the second core function *ccs_core_fastGraph*. The basic idea behind this function is to distribute the correlation computation among different parcels (blocks). For each parcel, *ccs_core_fastCoRR* calculates the correlation matrix between the parcel and the whole brain. Due to limitations in the memory space available to store the whole-brain correlation matrix, this parcel-specific correlation matrix must be compressed into a sparse representation by dropping all correlations below a certain threshold. Finally, we obtain a sparse adjacency matrix (binary or weighted) to represent the whole-brain graph by concatenating all parcel-specific sparse matrices. Figure 2 illustrates the algorithmic principle of *ccs_core_fastGraph* based upon the Yeo2011_N1000 Functional Atlas, including seven common neural networks [47]: the visual (Visual), somatomotor (SomMot), dorsal attention (DorsAttn), ventral attention (VentAttn), limbic (Limbic), frontoparietal (Control) and default (Default) networks. These two functions give the CCS the ability to map individual human functional connectomes on both super- and personal computers with very high extendibility of integrating novel algorithms (e.g., dynamical connectomics feature [75]).

5 H3: Connectome mining and knowledge discovery

How can we turn maps generated by the CCS into meaningful biological knowledge? Three major types of method

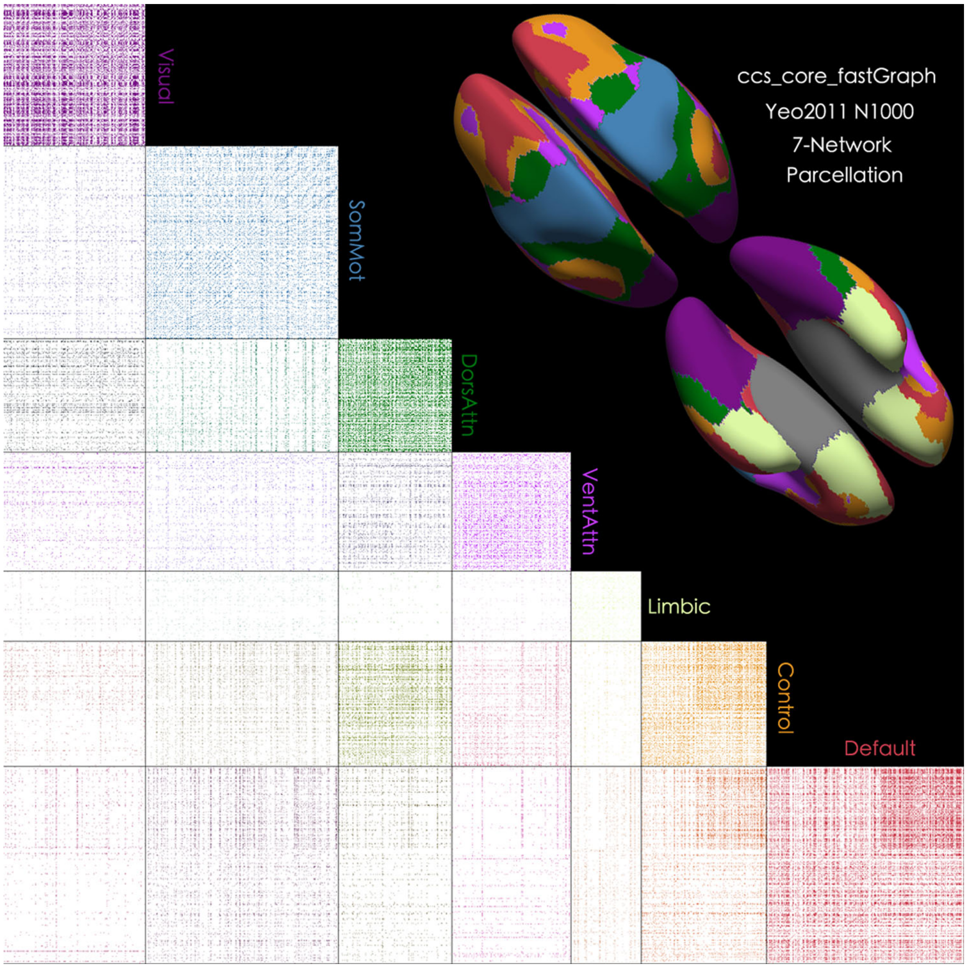

Fig. 2 Fast reconstruction of correlational brain graphs. A brain-inspired block matrix computation is proposed (ccs_core_fastGraph), wherein a large-scale computation is distributed to different brain modules derived from the Yeo2011_N1000 Functional Atlas. Seven common networks regarding the intrinsic functional architecture of the human brain are visualized onto the cortical surfaces in both dorsal and ventral views: the visual (Visual), somatomotor (SomMot), dorsal attention (DorsAttn), ventral attention (VentAttn), limbic (Limbic), frontoparietal (Control) and default mode (Default) networks. A final graph matrix of the full brain connectome can be delivered to 28 block matrices with different sizes where information compression (e.g., thresholding the correlation coefficient of a time series) and graph-based processing can be easily implemented, even on a personal computer

can be applied to mine connectome data and uncover knowledge, including regression/clustering, machine learning and information visualization. The first two methods have been widely used in the neuroimaging field [76], but the third has not been fully recognized by the field and has rarely been implemented in neuroimaging pipelines. The CCS implements different variants of these methods. Classic linear regression models are easily carried out in various neuroimaging toolboxes (such as SPM, FSL, AFNI, FreeSurfer). The CCS provides a set of script templates to make an easy and efficient translation between different toolboxes. At the next level of data mining, nonlinear and machine-learning algorithms are embedded into the last hierarchy of the CCS by combining MATLAB and R, where all resources developed in bioinformatics for genomics data mining and knowledge discovery can be reused in brain connectomics (e.g., pattern matching algorithms [77]). Finally, the CCS offers a toolkit of visualization to produce informative and easy-to-read figures for knowledge discovery and hypothesis generation. In the following sections, we will describe two representative algorithms; gRAICAR [78, 79] and test–retest reliability [60].

gRAICAR is an unsupervised machine-learning method used to mine variability in brain connectomes. One advantage of gRAICAR is that it characterizes inter-subject variability of connectomics metrics without a priori assumptions on subject groupings. This is a highly desirable feature for neuroimaging data mining, because the inter-subject variability in brain connectomes is not well mapped to our current categorizations of behavior and clinical symptoms. For instance, our current diagnosis of mental disorders depends solely on behavioral symptoms, whereas the biological mechanisms underlying different categories of mental disorders are not clearly bounded. This has been considered a primary cause of the lack of progress in translating brain connectome findings to clinical practice in psychiatry [80]. It is therefore necessary to establish the inter-subject variability in the connectome, based on which novel hypotheses on subject categorization can be proposed.

In functional connectomics, gRAICAR first decomposes the variation of functional images for each subject into a number of components using independent component analysis. The resultant components are spatial maps, each attributed to a mixing time course that represents the variation of the corresponding component map over time. gRAICAR then computes a similarity matrix across all component maps from all subjects (assuming N subjects). Based on the similarity matrix, gRAICAR identifies a component that has the highest similarity with $N - 1$ components, each from a different subject. These N components are thus grouped into an aligned component that reflects an intrinsic connectivity network. This procedure is repeated until all components are assigned to an aligned component. Each aligned component reflects a set of component maps from different subjects, which exhibit similar spatial distributions. An inter-subject similarity matrix is then constructed to depict the inter-subject relationship reflected by each aligned component. After learning the inter-subject relationship carried in the intrinsic connectivity networks, one can propose hypotheses based on the inter-subject relationship. These hypotheses can be further examined and interpreted using behavioral and clinical characteristics. Applications of gRAICAR have been demonstrated in a life span trajectory study of the functional connectome [81] and in a novel classification of early-onset schizophrenia patients [82].

Another example of CCS use for valuing functional connectomics is its component of test–retest reliability and reproducibility. High test–retest reliability is required to ensure the temporal stability of a metric and to allow for the distinguishing of different individuals [60]. This is a requirement for developing a biomarker of the application of functional connectomics, such as mapping growth charts of human brain function [83, 84]. The reliability of a measure is the upper bound of the correlation between the measure and another measure [60]. Therefore, beyond developing a biomarker, estimations of the test–retest reliability of functional connectomics are valuable for providing a reference regarding how strongly the variables affect the observed results and help to guide explanations of the findings of both normal and abnormal brains. The CCS employs linear mixed models to extract both intra- and inter-individual variability and develops a module on test–retest reliability measured by the intra-class correlation coefficients (ICC). This module not only provides various functions for evaluating the reliability and reproducibility of findings, but also accelerates the establishment of data analysis standards for functional connectomics based upon big test–retest neuroimaging data such as the data from the Consortium for Reliability and Reproducibility (CoRR [51]).

6 Case demonstration: normative trajectories of brain development

!The human brain experiences changes of its structure and function across the life span of an individual [67, 81, 84–90]. These changes can be altered by genetic and environmental variables, likely leading to brain disorders at different stages of the life span [91]. A normative life span trajectory of the human brain thus becomes extremely useful for early detection and diagnosis of these brain disorders, as well as for monitoring the effects of treatment effects. Here, we present a case usage of the CCS on delineation of normative trajectories of large-scale functional networks' morphology over the life span.

A total of 418 participants from the NKI–Rockland life span Sample (NKI-RS) [92] are included in the present analysis. There are 32 subjects who were excluded from subsequent analyses due to clinical diagnoses as defined by DSM-IV or ICD10 or due to incompleteness of the multimodal imaging datasets. Each participant has one anatomical image, one diffusion structural image and three resting-state functional images. The rfMRI images were obtained with three different scanning sequence settings to sample human brain function across several spatial and temporal scales: (1) std2500 (temporal resolution = 2,500 ms, spatial resolution = 3 mm, scan duration = 5 min); (2) mx1400 (temporal resolution = 1,400 ms, spatial resolution = 2 mm, scan duration = 10 min); and (3) mx645 (temporal resolution = 645 ms, spatial resolution = 3 mm and scan duration = 10 min). All of these data are publicly accessible via the FCP/INDI Web site (http://fcon_1000.projects.nitrc.org/indi/enhanced/index.html).

All image data were preprocessed in the CCS pipeline, which took approximately 15,000 CPU h in the Dell Blade Cluster System at the Institute of Psychology, Chinese Academy of Sciences. The QCP in the CCS filtered out 64 subjects by excluding low-quality multimodal imaging datasets, which met at least one of the following criteria: (1) failed in visual inspection on anatomical images and surfaces; (2) meanFD > 0.2 mm; (3) maxTrans > 3 mm; (4) maxRot $> 3°$; and (5) mcBBR > 0.6. This gives a final life span (6–85 years) multimodal neuroimaging sample of 316 healthy participants. These preprocessed images were then passed through the CCS pipeline to produce the various individual connectomics metrics (see H2). All outcome data from H1 and H2 will be made free to the public soon via a neuroimaging data-sharing platform developed by our team.

To demonstrate the feasibility of the CCS at discovering multimodal brain imaging data, we chose to model the life

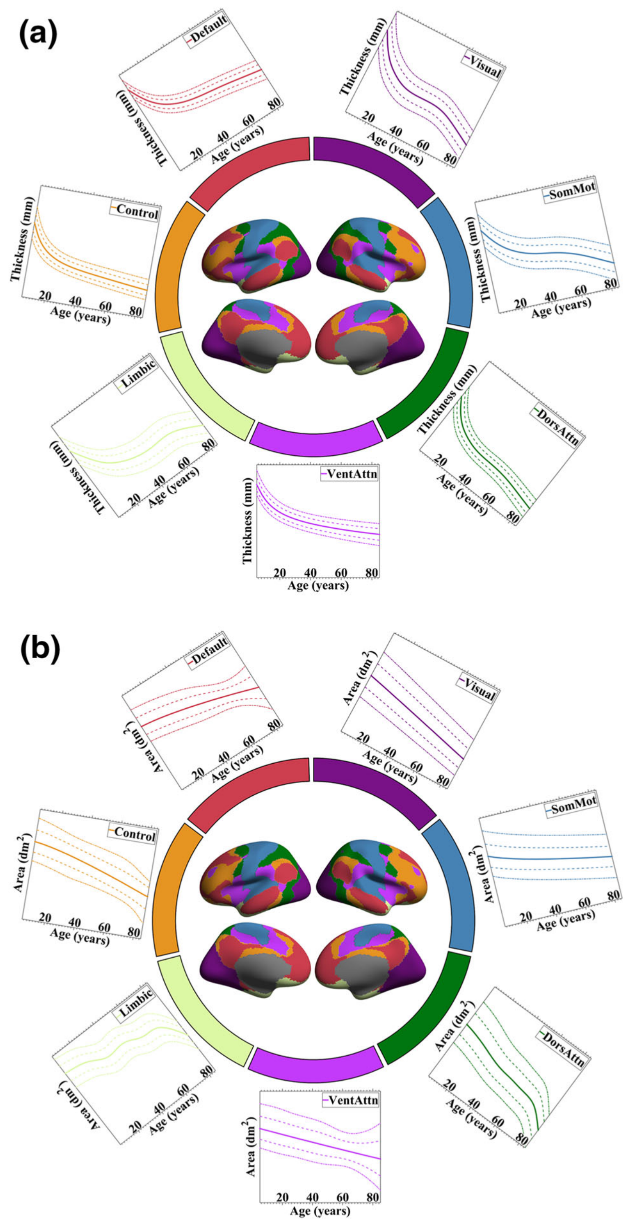

Fig. 3 Lifespan trajectories of seven neural networks. The network and lifespan developmental trajectories of (**a**) cortical thickness and (**b**) surface area are depicted in different colors. Centiles of 10 %, 25 %, 50 %, 75 % and 90 % are estimated for these curves across the human life span (5–90 years of age)

span trajectories of two morphological measures (cortical thickness and surface area) across the common large-scale neural networks in the Yeo2011_N1000 Functional Atlas [47]. Life span trajectory percentiles were estimated with generalized additive models for location, scale and shape (GAMLSS) [93], which extend the least mean square (LMS) method of the World Health Organization (WHO) standard methodology for estimation of growth charts of height and weight [94]. This is a semi-parametric statistical-modeling technique that allows estimation of age-specific percentiles and z scores. Models were fit in accordance with WHO methodology using cubic smoothing splines. Model selection was based on a two-step local maximum likelihood [95].

The 90 %, 75 %, 50 %, 25 % and 10 % centiles of the lifespan trajectories of cortical thickness (Fig. 3a) and surface area (Fig. 3b) are depicted for each of the seven neural networks. This visualization aids in characterization of the human brain connectome and its life span dynamics: (1) thickness and area are two distinct morphological metrics with regard to different life span trajectories; (2) all neural networks develop quickly but plateau with different speeds and at different critical ages; (3) these neural networks have different durations of plateau (maturation process); and (4) aging processes occur on a slow timescale but with different speeds. These observations will be comprehensively investigated in our future studies.

7 Conclusion

CCS is a computational pipeline useful for preprocessing multimodal neuroimaging MRI data, mapping features of individual brain connectomes, performing map mining and group-level statistics as well as uncovering novel knowledge. We hope a great service to the connectomics field by sharing and developing CCS with the community.

Acknowledgments This work was partially supported by the National Basic Research Program (973) of China (2015CB351702), the National Natural Science Foundation of China (81220108014, 81471740, 81201153, 81171409, and 81270023), the Key Research Program (KSZD-EW-TZ-002) and the Hundred Talents Program of the Chinese Academy of Sciences. Dr. Xiu-Xia Xing acknowledges the Beijing Higher Education Young Elite Teacher Project (No. YETP1593). Dr. Zhi Yang acknowledges the Foundation of Beijing Key Laboratory of Mental Disorders (2014JSJB03) and the Outstanding Young Researcher Award from Institute of Psychology, Chinese Academy of Sciences (Y4CX062008). We thank all members of the Laboratory for Functional Connectome and Development, Institute of Psychology at CAS and the attendees of the first CCS education course for their helpful feedback and suggestions for the improvement of the CCS.

Conflict of interest The authors declare that they have no conflicts of interest.

References

1. Sporns O, Tononi G, Kötter R (2005) The human connectome: a structural description of the human brain. PLoS Comput Biol 1:e42
2. Bullmore ET, Bassett DS (2011) Brain graphs: graphical models of the human brain connectome. Annu Rev Clin Psychol 7:113–140
3. Bullmore E, Sporns O (2012) The economy of brain network organization. Nat Rev Neurosci 13:336–349
4. Bullmore E, Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci 10:186–198
5. Breakspear M, Jirsa V, Deco G (2010) Computational models of the brain: from structure to function. Neuroimage 52:727–730

6. Deco G, Jirsa VK, McIntosh AR (2011) Emerging concepts for the dynamical organization of resting-state activity in the brain. Nat Rev Neurosci 12:43–56
7. Deco G, Jirsa VK, McIntosh AR (2013) Resting brains never rest: computational insights into potential cognitive architectures. Trends Neurosci 36:268–274
8. Song HF, Kennedy H, Wang XJ (2014) Spatial embedding of structural similarity in the cerebral cortex. Proc Natl Acad Sci USA 111:16580–16585
9. Chen Y, Wang S, Hilgetag CC et al (2013) Trade-off between multiple constraints enables simultaneous formation of modules and hubs in neural systems. PLoS Comput Biol 9:e1002937
10. Biswal BB, Mennes M, Zuo XN et al (2010) Toward discovery science of human brain function. Proc Natl Acad Sci USA 107:4734–4739
11. Seung HS (2011) Neuroscience: towards functional connectomics. Nature 471:170–172
12. Alivisatos AP, Chun M, Church GM et al (2012) The brain activity map project and the challenge of functional connectomics. Neuron 74:970–974
13. Smith SM, Vidaurre D, Beckmann CF et al (2013) Functional connectomics from resting-state fMRI. Trends Cogn Sci 17:666–682
14. Lander ES, Linton LM, Birren B et al (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921
15. Venter JC, Adams MD, Myers EW et al (2001) The sequence of the human genome. Science 291:1304–1351
16. 1000 Genomes Project Consortium (2010) A map of human genome variation from population-scale sequencing. Nature 467:1061–1073
17. 1000 Genomes Project Consortium (2012) An integrated map of genetic variation from 1,092 human genomes. Nature 491:56–65
18. Van Essen DC, Smith SM, Barch DM et al (2013) The WU-Minn human connectome project: an overview. Neuroimage 80:62–79
19. Schadt EE, Linderman MD, Sorenson J et al (2010) Computational solutions to large-scale data management and analysis. Nat Rev Genet 11:647–657
20. Berger B, Peng J, Singh M (2013) Computational solutions for omics data. Nat Rev Genet 14:333–346
21. Turk-Browne NB (2013) Functional interactions as big data in the human brain. Science 342:580–584
22. Marcus DS, Olsen TR, Ramaratnam M et al (2007) The Extensible Neuroimaging Archive Toolkit: an informatics platform for managing, exploring, and sharing neuroimaging data. Neuroinformatics 5:11–34
23. Scott A, Courtney W, Wood D et al (2011) COINS: an innovative informatics and neuroimaging tool suite built for large heterogeneous datasets. Front Neuroinform 5:33
24. Craddock RC, Jbabdi S, Yan CG et al (2013) Imaging human connectomes at the macroscale. Nat Methods 10:524–539
25. Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage 9:179–194
26. Fischl B, Sereno MI, Dale AM (1999) Cortical surface-based analysis. II. Inflation, flattening, and a surface-based coordinate system. Neuroimage 9:195–207
27. Segonne F, Dale AM, Busa E et al (2004) A hybrid approach to the skull stripping problem in MRI. Neuroimage 22:1060–1075
28. Segonne F, Pacheco J, Fischl B (2007) Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. IEEE Trans Med Imaging 26:518–529
29. Xing XX, Zhou YL, Adelstein JS et al (2011) PDE-based spatial smoothing: a practical demonstration of impacts on MRI brain extraction, tissue segmentation and registration. Magn Reson Imaging 29:731–738
30. Zuo XN, Xing XX (2011) Effects of non-local diffusion on structural MRI preprocessing and default network mapping: statistical comparisons with isotropic/anisotropic diffusion. PLoS One 6:e26703
31. Eskildsen SF, Coupé P, Fonov V et al (2012) BEaST: brain extraction based on nonlocal segmentation technique. Neuroimage 59:2362–2373
32. Klein A, Andersson J, Ardekani BA et al (2009) Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. Neuroimage 46:786–802
33. Behrens TE, Woolrich MW, Jenkinson M et al (2003) Characterization and propagation of uncertainty in diffusion-weighted MR imaging. Magn Reson Med 50:1077–1088
34. Andersson JL, Skare S (2002) A model-based method for retrospective correction of geometric distortions in diffusion-weighted EPI. Neuroimage 16:177–199
35. Chang LC, Jones DK, Pierpaoli C (2005) RESTORE: robust estimation of tensors by outlier rejection. Magn Reson Med 53:1088–1095
36. Beaulieu C, Allen PS (1994) Determinants of anisotropic water diffusion in nerves. Magn Reson Med 31:394–400
37. Mori S, van Zijl PC (2002) Fiber tracking: principles and strategies—a technical review. NMR Biomed 15:468–480
38. Greve DN, Fischl B (2009) Accurate and robust brain image alignment using boundary-based registration. Neuroimage 48:63–72
39. Behrens TE, Berg HJ, Jbabdi S et al (2007) Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? Neuroimage 34:144–155
40. Taylor PA, Saad ZS (2013) FATCAT: (an efficient) functional and tractographic connectivity analysis toolbox. Brain Connect 3:523–535
41. Saad ZS, Reynolds RC, Jo HJ et al (2013) Correcting brain-wide correlation differences in resting-state FMRI. Brain Connect 3:339–352
42. Power JD, Mitra A, Laumann TO et al (2014) Methods to detect, characterize, and remove motion artifact in resting state fMRI. Neuroimage 84:320–341
43. Carp J (2013) Optimizing the order of operations for movement scrubbing: comment on power. Neuroimage 76:436–438
44. Yan CG, Cheung B, Kelly C et al (2013) A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics. Neuroimage 76:183–201
45. Satterthwaite TD, Elliott MA, Gerraty RT et al (2013) An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. Neuroimage 64:240–256
46. Jo HJ, Saad ZS, Simmons WK et al (2010) Mapping sources of correlation in resting state FMRI, with artifact detection and removal. Neuroimage 52:571–582
47. Yeo BT, Krienen FM, Sepulcre J et al (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol 106:1125–1165
48. Fox MD, Zhang D, Snyder AZ et al (2009) The global signal and observed anticorrelated resting state brain networks. J Neurophysiol 101:3270–3283
49. Murphy K, Birn RM, Handwerker DA et al (2009) The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? Neuroimage 44:893–905
50. Yan CG, Craddock RC, Zuo XN et al (2013) Standardizing the intrinsic brain: towards robust measurement of inter-individual variation in 1000 functional connectomes. Neuroimage 80:246–262
51. Zuo XN, Anderson JS, Bellec P et al (2014) An open science resource for establishing reliability and reproducibility in functional connectomics. Sci Data 1:140049

52. He Y, Chen ZJ, Evans AC (2007) Small-world anatomical networks in the human brain revealed by cortical thickness from MRI. Cereb Cortex 17:2407–2419
53. Mechelli A, Friston KJ, Frackowiak RS et al (2005) Structural covariance in the human cortex. J Neurosci 25:8303–8310
54. Evans AC (2013) Networks of anatomical covariance. Neuroimage 80:489–504
55. Alexander-Bloch A, Giedd JN, Bullmore E (2013) Imaging structural co-variance between human brain regions. Nat Rev Neurosci 14:322–336
56. Rubinov M, Sporns O (2010) Complex network measures of brain connectivity: uses and interpretations. Neuroimage 52:1059–1069
57. Hagmann P, Cammoun L, Gigandet X et al (2008) Mapping the structural core of human cerebral cortex. PLoS Biol 6:e159
58. Gong G, Rosa-Neto P, Carbonell F et al (2009) Age- and gender-related differences in the cortical anatomical network. J Neurosci 29:15684–15693
59. van den Heuvel MP, Sporns O (2011) Rich-club organization of the human connectome. J Neurosci 31:15775–15786
60. Zuo XN, Xing XX (2014) Test–retest reliabilities of resting-state FMRI measurements in human brain functional connectomics: a systems neuroscience perspective. Neurosci Biobehav Rev 45:100–118
61. Zuo XN, Xu T, Jiang L et al (2013) Toward reliable characterization of functional homogeneity in the human brain: preprocessing, scan duration, imaging resolution and computational space. Neuroimage 65:374–386
62. Jiang L, Xu T, He Y et al (2014) Toward neurobiological characterization of functional homogeneity in the human cortex: regional variation, morphological association and functional covariance network organization. Brain Struct Funct. doi:10.1007/s00429-014-0795-8
63. Zang YF, He Y, Zhu CZ et al (2007) Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI. Brain Dev 29:83–91
64. Zou QH, Zhu CZ, Yang Y et al (2008) An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: fractional ALFF. J Neurosci Methods 172:137–141
65. Zuo XN, Di Martino A, Kelly C et al (2010) The oscillating brain: complex and reliable. Neuroimage 49:1432–1445
66. Zang Y, Jiang T, Lu Y et al (2004) Regional homogeneity approach to fMRI data analysis. Neuroimage 22:394–400
67. Zuo XN, Kelly C, Di Martino A et al (2010) Growing together and growing apart: regional and sex differences in the lifespan developmental trajectories of functional homotopy. J Neurosci 30:15034–15043
68. Biswal B, Yetkin FZ, Haughton VM et al (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn Reson Med 34:537–541
69. Greicius MD, Krasnow B, Reiss AL et al (2003) Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc Natl Acad Sci USA 100:253–258
70. Zuo XN, Kelly C, Adelstein JS et al (2010) Reliable intrinsic connectivity networks: test–retest evaluation using ICA and dual regression approach. Neuroimage 49:2163–2177
71. Zuo XN, Ehmke R, Mennes M et al (2012) Network centrality in the human functional connectome. Cereb Cortex 22:1862–1875
72. Destrieux C, Fischl B, Dale A et al (2010) Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage 53:1–15
73. Wu X, Xu L, Yao L (2014) Big data analysis of the human's functional interactions based on fMRI. Chin Sci Bull 59:5059–5065
74. Loewe K, Grueschow M, Stoppel CM et al (2014) Fast construction of voxel-level functional connectivity graphs. BMC Neurosci 15:78
75. Liao W, Wu GR, Xu Q et al (2014) DynamicBC: a MATLAB toolbox for dynamic brain connectome analysis. Brain Connect. doi:10.1089/brain.2014.0253
76. Bowman FD (2014) Brain imaging analysis. Annu Rev Stat Appl 1:61–85
77. Xue SW, Weng XC, He S et al (2013) Similarity representation of pattern-information fMRI. Chin Sci Bull 58:1236–1242
78. Yang Z, Zuo XN, Wang P et al (2012) Generalized RAICAR: discover homogeneous subject (sub)groups by reproducibility of their intrinsic connectivity networks. Neuroimage 63:403–414
79. Yang Z, LaConte S, Weng X et al (2008) Ranking and averaging independent component analysis by reproducibility (RAICAR). Neuroimage 63:403–414
80. Kapur S, Phillips AG, Insel TR (2013) Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? Mol Psychiatry 17:1174–1179
81. Yang Z, Chang C, Xu T et al (2012) Connectivity trajectory across lifespan differentiates the precuneus from the default network. Neuroimage 89:45–56
82. Yang Z, Xu Y, Xu T et al (2014) Brain network informed subject community detection in early-onset schizophrenia. Sci Rep 4:5549
83. Castellanos FX, Di Martino A, Craddock RC et al (2013) Clinical applications of the functional connectome. Neuroimage 80:527–540
84. Dosenbach NU, Nardos B, Cohen AL et al (2010) Prediction of individual brain maturity using fMRI. Science 329:1358–1361
85. Collin G, van den Heuvel MP (2013) The ontogeny of the human connectome: development and dynamic changes of brain connectivity across the life span. Neuroscientist 19:616–628
86. Cao M, Wang JH, Dai ZJ et al (2014) Topological organization of the human brain functional connectome across the lifespan. Dev Cogn Neurosci 7:76–93
87. Betzel RF, Byrge L, He Y et al (2014) Changes in structural and functional connectivity among resting-state networks across the human lifespan. Neuroimage 102:345–357
88. Chan MY, Park DC, Savalia NK et al (2014) Decreased segregation of brain systems across the healthy adult lifespan. Proc Natl Acad Sci USA 111:E4997–E5006
89. Yeatman JD, Wandell BA, Mezer AA (2014) Lifespan maturation and degeneration of human brain white matter. Nat Commun 5:4932
90. Gutchess A (2014) Plasticity of the aging brain: new directions in cognitive neuroscience. Science 346:579–582
91. Di Martino A, Fair DA, Kelly C et al (2014) Unraveling the miswired connectome: a developmental perspective. Neuron 83:1335–1353
92. Nooner KB, Colcombe SJ, Tobe RH et al (2012) The NKI-Rockland sample: a model for accelerating the pace of discovery science in psychiatry. Front Neurosci 6:152
93. Rigby RA, Stasinopoulos DM (2005) Generalized additive models for location, scale and shape. J R Stat Soc Ser C Appl Stat 54:507–554
94. Multicentre Growth Reference Study Group WHO (2009) WHO Child Growth Standards: growth velocity based on weight, length and head circumference: methods and development. World Health Organization, Geneva
95. Rigby RA, Stasinopoulos DM (2013) Automatic smoothing parameter selection in GAMLSS with an application to centile estimation. Stat Methods Med Res 23:318–332