



Computational neuroscience

## Mapping the human connectome at multiple scales with diffusion spectrum MRI

Leila Cammoun<sup>a,\*</sup>, Xavier Gigandet<sup>a</sup>, Djalel Meskaldji<sup>a</sup>, Jean Philippe Thiran<sup>a</sup>,  
Olaf Sporns<sup>b</sup>, Kim Q. Do<sup>c</sup>, Philippe Maeder<sup>d</sup>, Reto Meuli<sup>d</sup>, Patric Hagmann<sup>a,d</sup>

<sup>a</sup> Signal Processing Laboratory (LTS5), Ecole Polytechnique Fédérale de Lausanne, Switzerland

<sup>b</sup> Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, United States

<sup>c</sup> Center of Psychiatric Neuroscience, University Hospital Center and University of Lausanne (CHUV), Lausanne, Switzerland

<sup>d</sup> Department of Radiology, University Hospital Center and University of Lausanne (CHUV), Lausanne, Switzerland

### ARTICLE INFO

#### Article history:

Received 17 March 2010

Received in revised form 24 August 2011

Accepted 29 September 2011

#### Keywords:

Brain connectivity  
Structural networks  
Diffusion MRI  
Connectomics  
Connection matrix  
Multi-scale

### ABSTRACT

The global structural connectivity of the brain, the human connectome, is now accessible at millimeter scale with the use of MRI. In this paper, we describe an approach to map the connectome by constructing normalized whole-brain structural connection matrices derived from diffusion MRI tractography at 5 different scales. Using a template-based approach to match cortical landmarks of different subjects, we propose a robust method that allows (a) the selection of identical cortical regions of interest of desired size and location in different subjects with identification of the associated fiber tracts (b) straightforward construction and interpretation of anatomically organized whole-brain connection matrices and (c) statistical inter-subject comparison of brain connectivity at various scales. The fully automated post-processing steps necessary to build such matrices are detailed in this paper. Extensive validation tests are performed to assess the reproducibility of the method in a group of 5 healthy subjects and its reliability is as well considerably discussed in a group of 20 healthy subjects.

© 2011 Elsevier B.V. All rights reserved.

### 1. Introduction

The study of neuronal connections in the brain has been a difficult and demanding. Our current knowledge of brain connectivity is largely based on the study of the relationship between symptoms and lesions as well as on post-mortem dissections of large fiber tracts. The former approach was pioneered by Broca (1861) and Wernicke (1906), and the latter by Gall and Spurzheim (1810–1819) and others. More recently, however, great strides have been made in chemical tracing methods in the macaque (Schmahmann and Pandya, 2006) as well as in humans (Clarke et al., 1999; Di Virgilio et al., 1999; Stephan et al., 2008; Zaidel et al., 1995), which have allowed the identification of not only gross fiber tracts but also individual white matter connections. These efforts have resulted in the definitive mapping of a few tens of connections in humans and several hundreds in the macaque. Although such tracing studies are immensely useful and of high-resolution they are also very limited since they are confined to post-mortem material and each study is limited to a few connections only. New high throughput techniques are needed. Important advances have been made with the advent of diffusion MRI tractography, which circumvents the drawbacks mentioned above by allowing not only

in vivo (Conturo et al., 1999; Hagmann et al., 2003; Mori et al., 1999; Wedeen, 1996) but also post-mortem imaging (Schmahmann et al., 2007) of a large number of fiber bundles, this, however at the cost of lower resolution. These techniques have spurred many studies related to normal or pathologic neuro-anatomy. More recently it became clear that beyond the aim of characterizing individual fiber bundles, the connectivity profile of the entire brain is of highest importance in neuroscience. Following the pioneering work based on chemical tracing of (Felleman and Van Essen, 1991) and others (Hilgetag and Kaiser, 2004; Hilgetag et al., 2000; Sporns and Zwi, 2004), similar connection matrices have been built from MRI tractography, either by constructing large-scale networks of 1000 nodes (Hagmann et al., 2007) or more anatomically based connection matrices (Gong et al., 2009, 2008; Iturria-Medina et al., 2007, 2008; Li et al., 2009; Thottakara et al., 2006). Diffusion-based connectivity has also been used in some studies (Behrens and Johansen-Berg, 2005; Klein et al., 2007; Tomassini et al., 2007) to parcellate gray matter. In order to emphasize the importance of whole brain connectivity, the term connectome has been coined by our group as early as 2005 (Hagmann, 2005; Hagmann et al., 2010; Sporns, 2008; Sporns et al., 2005). It refers to the complete description of the structural connectivity of the brain. More recently our group showed that cortical areas involving the default mode network (Raichle and Snyder, 2007) correspond to highly connected hubs defining the core of structural connectivity (Hagmann et al., 2008). We also showed through computational modeling how low

\* Corresponding author. Tel.: +41 21 693 46 22.

E-mail address: [leila.cammoun@epfl.ch](mailto:leila.cammoun@epfl.ch) (L. Cammoun).

frequency BOLD oscillations can be predicted from structural connectivity, highlighting once again the fundamental relevance of connectomic approaches (Honey et al., 2009).

Given the increasing interest in such approaches as well as emerging questions about the optimal scale and optimal representation of such connection matrices (Bassett et al., 2010; Fornito et al., 2010; Zalesky et al., 2010), we take the opportunity to present in detail our approach to map the connectome at multiple scales and extensively test its reliability and reproducibility.

## 2. Materials and methods

### 2.1. Overview

Fig. 1 gives an overview of the methodology employed. Basically the processing pipeline was divided into two pathways. On one side, the cortical surface was extracted from a high-resolution T1-weighted image and subdivided into 66 anatomical parcels by matching the most important sulci using atlas-based segmentation. Each anatomical parcel was then subdivided into small cortical ROIs of equal area, 998 ROIs in total. Third, these 998 ROIs were successively grouped in order to have 5 sets of embedded cortical partitions with 66, 133, 241, 483 and 998 ROIs, respectively. These subdivisions and merging of ROIs necessary to obtain the 5 sets of embedded cortical partitions are done only once on the template used for the atlas based segmentation and then applied on the images to be analyzed. On the other side, we performed whole brain tractography, which results in millions of virtual fibers spread over the brain. The combination of these two procedures allowed the construction of connection matrices at multiple scales by computing the connection density between each pair of ROIs at each scale. The whole procedures described by the flow chart in Fig. 2 are detailed in the following paragraphs. This work is developed basically in Matlab (<http://www.mathworks.com>), and with the help of Freesurfer software (<http://surfer.nmr.mgh.harvard.edu>) and its related distributed Matlab functions among other home-made developments.

### 2.2. MRI acquisition

After written informed consent and in accordance with our institutional guidelines and ethics committee, we scanned twenty healthy right-handed male volunteers in the age range of 22–35 years. The imaging protocol was performed using a Siemens TimTrio 3T whole body MRI system (with a 32-channel receive head matrix coil). *Parallel imaging was used with an acceleration factor of 2 and a Grappa reconstruction. The sequence is based on a Stejskal–Tanner spin echo EPI model, with a 90° flip angle, a TR of 6000 ms and a TE of 136 ms, and a maximal b-value of 8000 s/mm<sup>2</sup>. Q-space was sampled over a hemisphere in a cubic lattice with 258 points by varying the diffusion gradient intensity and direction such that  $\mathbf{q} = a\mathbf{q}_x + b\mathbf{q}_y + c\mathbf{q}_z$ , with  $a, b, c$  integers and  $\sqrt{a^2 + b^2 + c^2} \leq 5$ ,  $\mathbf{q}_x, \mathbf{q}_y, \mathbf{q}_z$  denoting the unit diffusion sensitizing gradient vectors in the three respective coordinate directions. The axial field of view was set to 212 by 212 mm and the size of the acquisition matrix was 96 by 96, yielding an in-plane resolution of 2.21 by 2.21 mm. 34 slices of 3 mm thickness were acquired yielding an acquisition time of 26 min. The reconstruction of the data was done according to the standard Diffusion Spectrum Imaging (DSI) protocol (Wedeen et al., 2005). By taking the Fourier transform of the modulus of the q-space signal in every voxel, the resulting probability density function was projected radially yielding an Orientation Distribution Function (ODF) in every voxel. In addition a high-resolution T1-weighted (MP-RAGE) MRI was*

acquired in a matrix of 256 × 256 × 128 voxels of isotropic 1 mm resolution.

We confirmed that head motion in the data sets was not significant. Eddy current distortions were minimized by maximizing parallel imaging using an acceleration factor of 3 and using compensated diffusion gradient. Post hoc eddy current correction was not felt to be useful.

### 2.3. Generation of normalized cortical ROIs by atlas based segmentation

For each subject, each ROI had to be placed in the same anatomical location such that connectivity could be compared locally. Every piece of this cortical puzzle should form a relatively compact surface of more or less constant size and should be as small as reasonable.

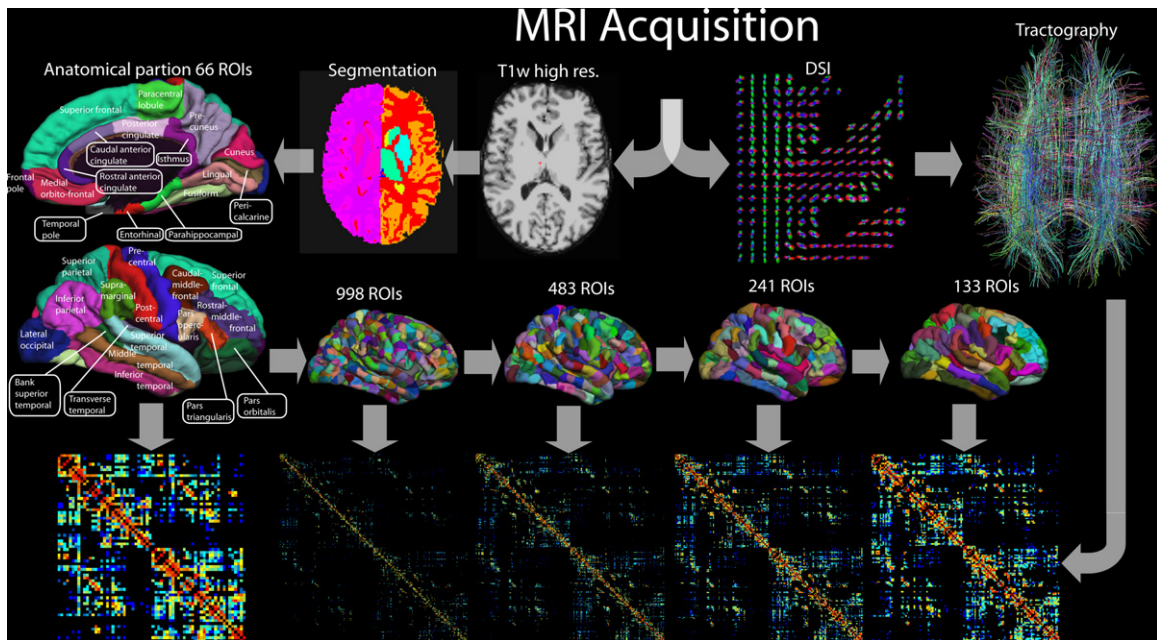
#### 2.3.1. Template creation

To fulfill these constraints our methodology relies on a template brain or atlas that was initially generated from 40 manually labeled brains where 66 anatomically typical cortical parcels were defined using curvature-based information (Desikan et al., 2006). These 66 cortical parcels available in FreeSurfer (<http://surfer.nmr.mgh.harvard.edu>) were, however, not small enough to obtain a high-resolution connection matrix. Ideally the ROIs should be small enough to approach the size of a voxel, but noise, partial volume effects, variance in the registration and brain anatomical variability would hamper reproducibility. With these issues in mind, we decided that ROIs should not be smaller than 1.5 cm<sup>2</sup> implying that the whole cortex would be divided approximately into 1000 cortical ROIs. Below is a description of how we obtained such a partition.

#### 2.3.2. Parcellation of the template cortical surface

To avoid any ambiguity in the text, we call the original gyral-based ROIs produced by Freesurfer ‘parcels’ and not ROIs. First, the number of ROIs per parcel, denoted by  $N_p$ , was calculated based on the relative surface area of each template parcel, as well as the desired total number of ROIs. This ensured that the variation of the ROI surface area between different parcels remains minimal. Then, for each parcel  $p$ , we perform a region growing from a randomly chosen point of  $p$  by aggregating the neighboring voxels, until it reached the desired surface area. This creates the first ROI of parcel  $p$ . Next, the other ROIs were generated from starting points located close to the already existing ROIs, until the parcel was fully covered with ROIs. In the second phase, we identified the  $N_p$  biggest ROIs of the parcel and computed their centers of gravity. We started a second growing process from these points, but this time with all ROIs growing simultaneously. Using this two-phase partitioning heuristic, we obtained ROIs that were compact and of similar size (Hagmann et al., 2007). As we forced the template ROIs to reach the same size, the resulting number of ROIs generated on the template cortex was actually 998 and not 1000. It is also important to emphasize that each parcel was subdivided individually, thus preventing the ROIs from overlapping on different parcels.

The original cortical partition is made up of  $N = 998$  ROIs. However the most suitable resolution between 998 and 66 ROIs actually depends on the application. Thus, we created a hierarchical decomposition between 66 and 998 ROIs by successive grouping. On the template brain, 2 or 3 neighboring ROIs at the 998 scales were manually grouped into one ROI to build a partition into 483 ROIs. This grouping operation was repeated several times until the 66 parcels were recovered. Using this heuristic, we ended up with 5 embedded cortical parcellations with  $N = \{998, 483, 241, 133, 66\}$ . In practice,



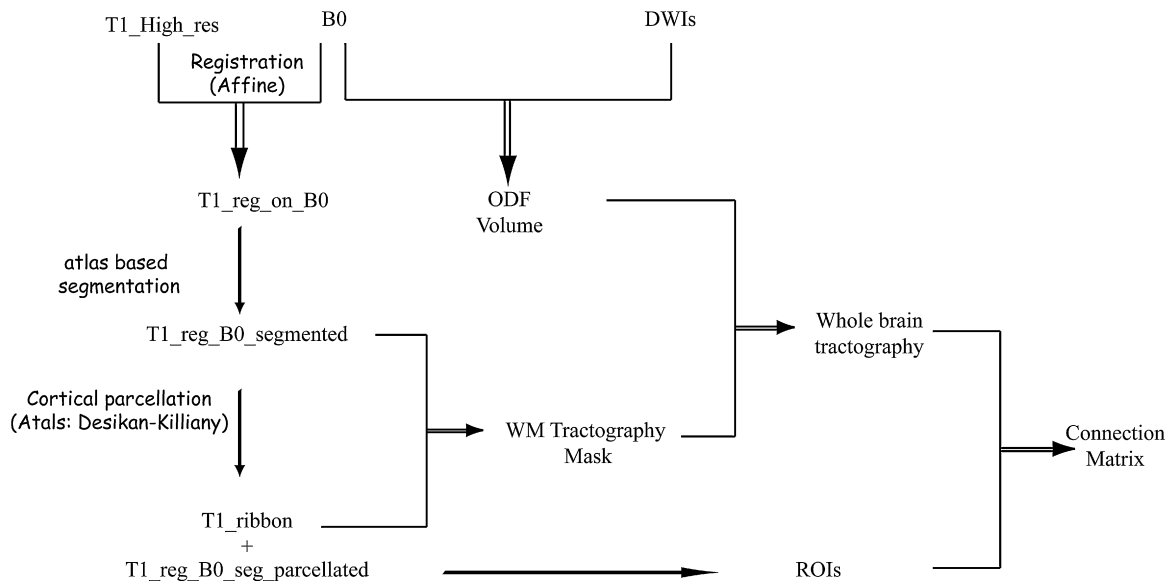
**Fig. 1.** Construction of normalized connection matrices. After the acquisition of the diffusion and high-resolution T1-weighted images: the ODF map is calculated from diffusion MRI and used to perform tractography. The T1-weighted image registered onto the space of diffusion data is segmented, and parcellated into 66 cortical parcels. Each parcel is subdivided into smaller ROIs of approximately identical surface such that the total number of ROIs is 998. The 998 cortical ROIs are regrouped iteratively into bigger ROIs such that in the end 5 scales of cortex parcellation are made available (66, 133, 241, 483, 998). The deep gray nuclei are added to cortical ROIs. Each set of ROIs is combined with the tractography result to create connection matrices of successive scales.

this means that every vertex of the mesh representing the cortex in the Freesurfer environment, had 5 labels identifying to what ROI it belongs for every resolution.

### 2.3.3. Cortical registration

Once the template cortex had been partitioned into a mosaic of ROIs at several resolutions, this parcellation could be transferred to the subject's cortex by means of cortical surface-based registration. The Freesurfer software (<http://surfer.nmr.mgh.harvard.edu>) offers a very reliable registration technique that is based on

curvature information, i.e., sulcus and gyrus (Desikan et al., 2006; Fischl et al., 2004). The cortical registration was carried out on every subject based on the high-resolution T1-weighted image. Through the registration of an individual parcel, by matching the sulcal boundaries, its mesh gets stretched and deformed but the neighborhood of individual vertices is preserved and as a matter of consequence of the ROIs lying in that parcel. It is worth mentioning that prior to this operation the T1-weighted image had been registered onto the space of the diffusion data with an affine registration technique (Van Leemput et al., 1999).



**Fig. 2.** Flow chart describing all the procedural steps. Anatomical T1 is registered on B0 with affine registration. Using Freesurfer software, the registered T1 is first segmented; sub-cortical nuclei are extracted in this step, as well as ventricles. A cortical parcellation is performed in a second step to obtain cortical parcels. The calculated transformation in this step is used with the constructed multi-resolution atlases (Fig. 1) to obtain the ROIs in respective resolutions. The WM mask is constructed using the cortical ribbon from which we remove non cortical ROIs as well as ventricles. Using the ODF from diffusion images, the tractography is performed in the WM mask. The final step is the construction of the multi-resolution connection matrices using the ROIs in each resolution and the virtual fibers.

## 2.4. Tractography

### 2.4.1. Building the tractography mask

In order to obtain consistent tractography results, the segmentation of the brain into white matter, cortex and deep cerebral nuclei is an essential step. In fact, the virtual fibers should be generated only in the white matter. To construct the tractography mask, the cortical surface obtained from FreeSurfer was first filled. Next, the ventricles and the deep gray nuclei (thalamus, pallidum, putamen, caudate nucleus accumbens, subthalamus) as well as the hippocampus and the amygdala were removed from the mask. The location of these structures was obtained by an atlas-based segmentation (Fischl et al., 2002), using the same software. Since all these structures except for the ventricles are relays for the cerebral fibers, they are furthermore considered as ROIs as well as the brain stem base, increasing the number of ROIs, respectively to {1015, 500, 258, 150, 83}.

### 2.4.2. White matter tractography

Tractography is a post-processing method based on the diffusion-weighted images. It constructs 3-dimensional curves of maximal diffusion coherence. These curves, called *virtual fibers*, are the estimates of the real white matter axonal bundle trajectories (Hagmann et al., 2004). We used a tractography algorithm specifically designed for DSI data to create a set of such fibers for the whole brain (Hagmann et al., 2007; Wedeen et al., 2008). At each voxel, a set of directions of maximum diffusion was defined as local maxima of the ODF. Then in each white matter voxel the same number of fibers for every direction of maximum diffusion were initiated. For example, in a voxel with 2 main directions, 4 fibers along each direction were initiated (total 8). The starting points were chosen at random within the voxel. Next, from each such point a growth of a fiber is started in two opposite directions with a fixed step of 1 mm. When entering a new voxel, the fiber growth continues along the direction of the vector of the maximum diffusion (in the new voxel) whose orientation is the closest to the current direction of the fiber. If this resulted in a change of direction sharper than  $45^\circ/0.5$  mm, the growth was stopped. The growth process of a valid fiber finished when both its ends left the white matter mask. In this paper we used about 3 million initialization points of which only about one half to two thirds of the resulting fibers connecting the cortical areas were retained.

## 2.5. Construction of the connection matrix

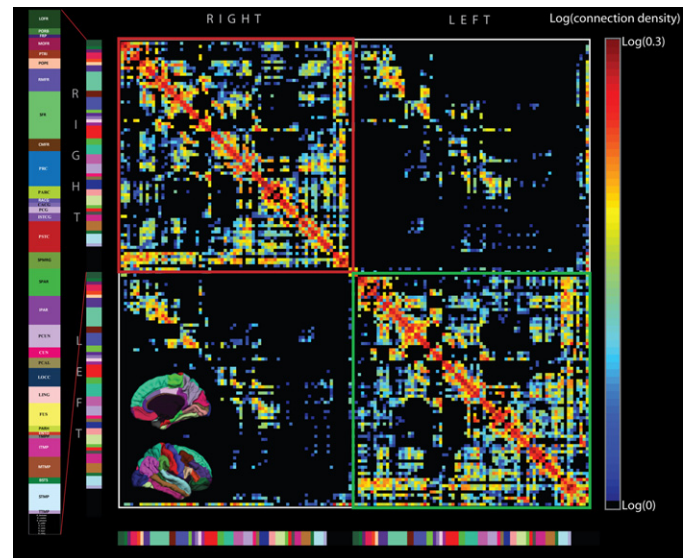
Considering the cortical parcellation and the white matter tractography described in the previous paragraphs, the fiber bundle  $B(i,j)$  connecting each pair of ROI  $i$  and  $j$  could be identified. The value of the connection matrix cell  $M(i,j)$  was the connection density between this pair of ROIs, defined as follows:

$$M(i,j) = \frac{2}{S(i) + S(j)} \sum_{f \in B(i,j)} \frac{1}{l(f)}$$

where  $S(\cdot)$  is the area of ROI  $(\cdot)$  and  $l(f)$  is the length of fiber  $f$  along its trajectory.

The correction term  $l(f)$  in the denominator is needed to eliminate the linear bias towards longer fibers introduced by the tractography algorithm (Hagmann et al., 2007) which uses each voxel in the white matter mask as a seed point.

To facilitate the visual interpretation of the connection matrix, we organized the matrix by taking into account as much as possible the ROI neighborhood. Inspired by brain development, we arranged the 33 parcels of each hemisphere in a fronto-caudal order. Then a linear order of the ROIs inside a parcel is intuitively the one corresponding to the direction of 'previous parcel to next parcel'.



**Fig. 3.** The average connection matrix of five healthy subjects at resolution 133, representing the fibers density between every pair of ROIs in logarithmic scale. The red and green squares correspond, respectively to right and left intra-hemispheric connections. Off diagonal squares represent the inter-hemispheric connections. X- and Y-axis color bars represent the parcels with the same color as template brain (see insert). The right side color bar indicates the logarithm of fiber density.

Since this organization was set up only once on a template brain, it remained identical when applied to different subjects.

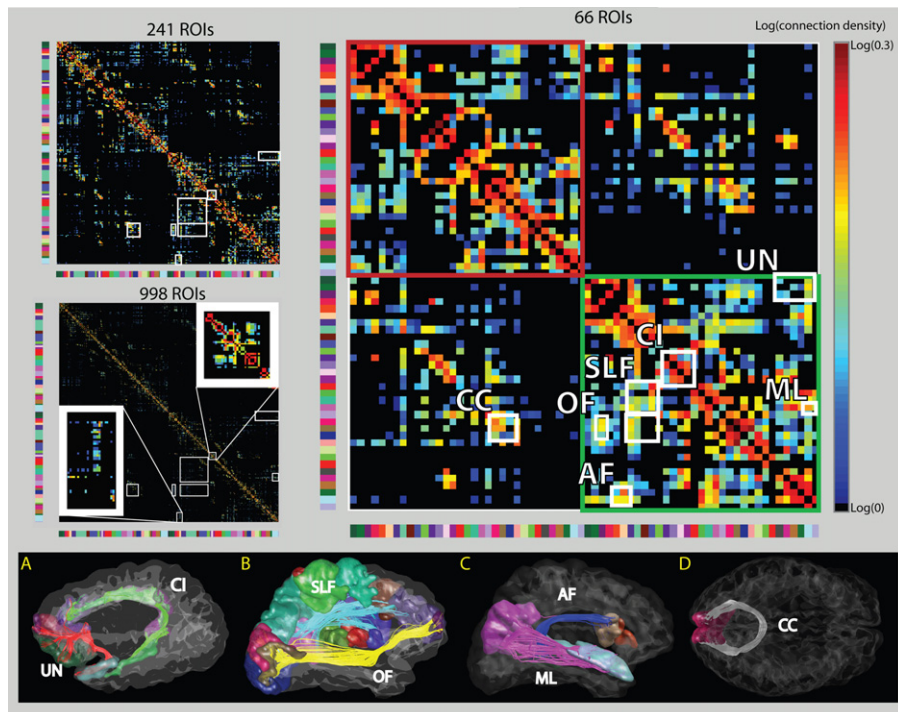
As a result, we obtained for each individual brain a set of 5 matrices (one per resolution) that were embedded in each other. The lowest resolution matrix entails as many entries as parcels in the standard FreeSurfer output, i.e., 66. The highest resolution matrix had 998 entries corresponding to 998 small ROIs (Fig. 1).

Fig. 3 represents the average connection matrix of five healthy subjects at resolution 133. The connection matrix was organized such that the upper left quadrant (with the red border) represents the connections in the right hemisphere and the lower right quadrant (with the green border) represents the left hemispheric connectivity. This matrix is symmetric since the measured connectivity is not oriented. The off-diagonal quadrants map the inter-hemispheric connections. The color bars at the left and bottom of the matrix help make the correspondence between the matrix entries and the 66 cortical parcels as displayed on the insert images. The color bar on the right codes the connection density seen in the matrix itself in logarithmic scale.

## 3. Results

### 3.1. Normalized connection matrices at multiple scales

Fig. 4 shows some of the connection matrices at different resolutions as they were computed in one subject. It is possible to identify known bundles from the connection matrix. In Fig. 4 we give several examples. We have selected groups of ROIs that are expected to correspond to language areas (Wernicke's and Broca's Area). The connections between these areas can easily be identified on the matrix and correspond to the arcuate fasciculus. The latter with the uncinate, the occipito-frontal, the middle longitudinal and the superior longitudinal fasciculi form long distance connections, which are accordingly far from the diagonal of the matrix. The cingulum bundle, which is made up mainly of sets of alternating short connections, is located close to the diagonal of the matrix. Another example is the pathways connecting the homotopic primary visual cortices, which are represented in the off-diagonal blocs. The occipito-frontal connection is represented



**Fig. 4.** Identification of different fiber bundles on the different connection matrices. The identified bundles are: cingulum (CI) and uncinate (UN) shown in panel A; superior longitudinal fasciculus (SLF) and occipito-frontal (OF) shown in panel B; arcuate fasciculus (AF) and middle longitudinal (ML) shown in panel C; and cuneus homotopic connection (CC) shown in panel D. All these bundles are also represented on other scale matrices.

by two squares in the matrix because not all the ROIs belonging to the frontal cortical area are neighbors in an arbitrary linear arrangement of the matrix entries. All these bundles are reported on higher resolution matrices where we could see these bundles spread out in the white boxes related to the corresponding ROIs.

As there is no *in vivo* brain connectivity gold standard, Fig. 4 is essentially considered as an illustration of some pathways in the connection matrix. Because we know that some of well-known bundles are large, imposing, connecting pairs of big regions, and some other are more subtle, connecting some small but known anatomical ROIs (from histology for instance), the multi-scale approach makes sense in this context. In fact, based on the current knowledge of the well-known pathways, the low-resolution matrices enable to map the connection between big bundles, connecting anatomical gyri. Nevertheless, the subdivision of the big parcels into smaller ROIs, normalized across subjects, enables to detail more and more one given bundle and see its ramification. We show in Fig. 5 the probability of existence of same bundles in three various resolutions through all the set of subjects. This was done by averaging the binarized matrices of all the subjects. We notice in these matrices, some red cells indicating the existence of bundles for all the subjects (Fig. 5A). In the same figure, to emphasize the node scale effect, we present for a single case a part of the arcuate fasciculus in various scales to show the power of the multi-resolution to identify the ROIs connecting this specific part of this well-known bundle Fig. 5C. We notice in Fig. 5B the existence of this bundle in the same cells in the resolution 83, some dark red cells still remain in resolution 258 but in the higher resolution we observe more variability among subjects.

### 3.2. Robustness

Since the ultimate aim of the methodology presented here is to study connectivity in groups of subjects, it is essential at this stage to evaluate its repeatability and robustness. The evaluation

has been done at the level of the ROI registration and of the connection matrix itself. We have proceeded in 3 steps: (1) we have processed 5 data sets twice and have compared both results. (2) We have scanned the same 5 subjects twice, processed the double data sets and compared both outputs. (3) We have scanned and processed 15 other healthy subjects and analyzed the connection matrices of the whole group of 20 subjects at all the different scales.

#### 3.2.1. Repeatability

##### (a) Same scans processed twice

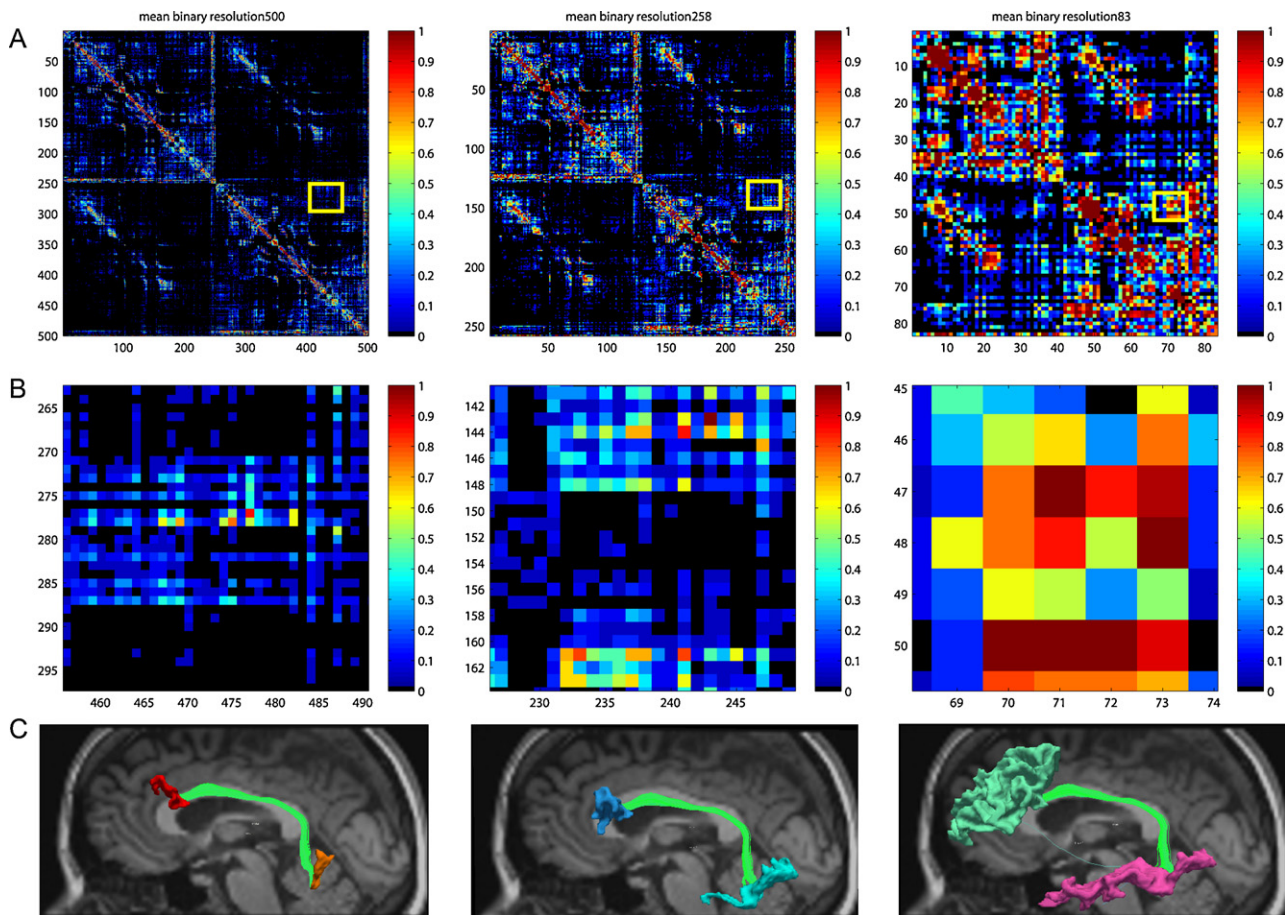
In order to evaluate the reproducibility of cortical ROI registration, and the complex pipeline used to achieve the connectome, we ran the whole processing twice on the same datasets for 5 subjects.

First, we evaluated the reproducibility of the method for the 5 scales by computing the Pearson correlation coefficient between the connection matrices resulting from the double processing of the same data set. Even if the matrices cells are dependents, the correlation coefficient is still a reasonable metric to compare individual connection matrices globally; furthermore the matrices of different individuals are independent.

If  $x$  and  $y$  are, respectively the elements of two matrices of size  $N$  by  $N$ , the Pearson correlation coefficient is calculated for the half of the matrices as they are symmetric and expressed as follows:

$$R(x, y) = \frac{\sum_{i=1:N} \sum_{j=i:N} (x_{ij} - \bar{x})(y_{ij} - \bar{y})}{\sqrt{\sum_{i=1:N} \sum_{j=i:N} (x_{ij} - \bar{x})^2 \sum_{i=1:N} \sum_{j=i:N} (y_{ij} - \bar{y})^2}} \quad \text{where}$$

$$\bar{x} = \frac{\sum_{i=1:N} \sum_{j=i:N} x_{ij}}{N(N+1)/2} \quad \text{and} \quad \bar{y} = \frac{\sum_{i=1:N} \sum_{j=i:N} y_{ij}}{N(N+1)/2}$$



**Fig. 5.** Multi-resolution and node effect: (A) the average of the binarized connection matrices of all the subjects in resolutions 500, 258 and 83. (B) Enlarged view of the connection matrices highlighted in yellow in A and representing the groups of ROIs that are expected to correspond to language areas (Wernicke's and Broca's area) and the arcuate fasciculus. The dark red cell corresponds to one meaning that for all the subjects the respective connection exists, and the light blue cell corresponds to only some subjects or one subject having this connection. (C) Illustration of the node scale effect on one single subject, on the part of arcuate fasciculus.

We see in Fig. 6, that the matrices resulting from the two independent processings for this sub group were highly correlated, but actually not equal to one since the whole processing pipeline contains some probabilistic processes. In fact, the first part of the processing using the Freesurfer software is not deterministic, as well as the initialization of the tractography algorithm. The correlation coefficients ranged between 0.977 and 0.998 depending on the resolution of the matrix. They decrease linearly when the number of ROIs increases, as would be expected since the small inaccuracies of the ROIs registration are more perceptible on small ROIs than in high resolution connection matrices. The  $p$ -values corresponding to the correlation coefficients are very low ( $p \ll 10e-5$ ).

To assess whether the performance of the method is homogeneous across bundle sizes (i.e., connection densities), a scatter plot of each pair of double connections is presented in Fig. 7A.

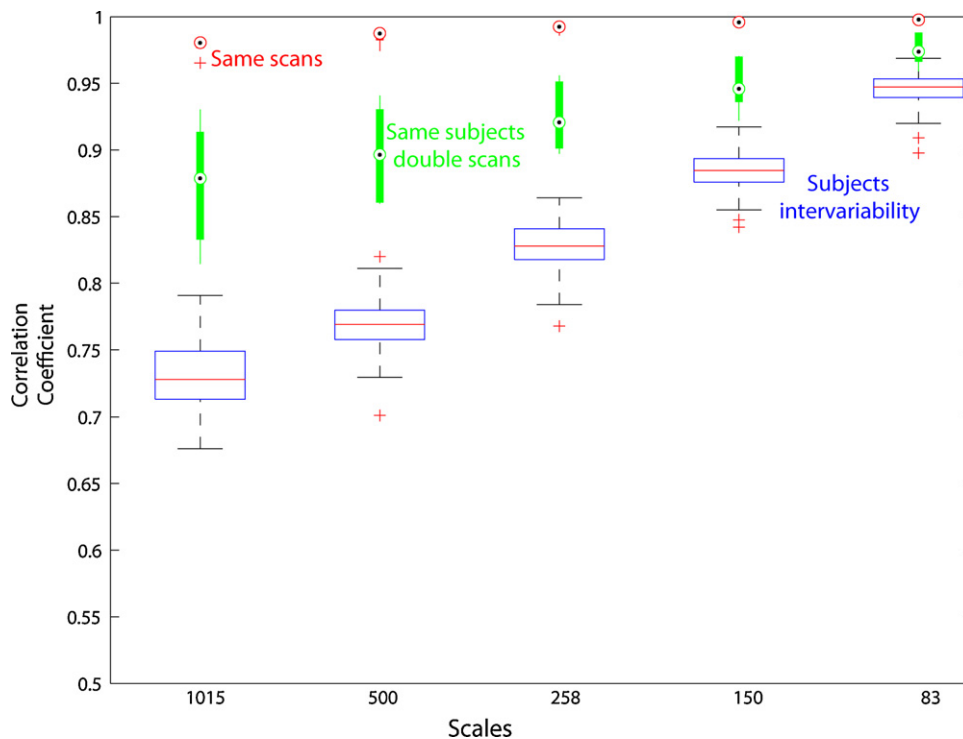
We can see in Fig. 7A that for all resolutions the repeated processings of the same subject data sets are highly correlated. The correlation coefficient of each double set is presented on the same figure with the same color as the data. However, a considerable difference in the scatter plot small indices (corresponding to small number of bundles) and large indices (corresponding to large number of bundles) is noteworthy. To further investigate the difference between the results of the reprocessing of the same subjects, the Bland–Altman plot (Bland and Altman, 1986) is presented in Fig. 7B. The interpretation of these two plots together points out that generally the big (i.e., dense) bundles are more reliably mapped than

the small ones at all resolutions. We also note that low resolutions yield the best correlation and the smallest difference between the double reprocessed data sets. This is probably due to the fact that at high-resolution, the small cortical ROIs introduce more noise because of the registration error. We can also see in this figure that the variance of observed value is smaller in low-resolution matrices. In fact the fiber numbers were averaged on bigger bundles linking bigger ROIs in the low-resolution matrices, reducing the noise of observations.

It is worth investigating whether some particular nodes of the matrix have more influence on the method stability, and particularly to assess this variation relative to bundle properties.

To assess this, the correlation coefficient matrix is calculated for every double set. An average correlation matrix is then obtained from the five original subject's correlation matrices. This is done for all resolutions. For a given resolution, e.g., the highest one, the correlation between the double subject processings is calculated for every cell connecting the original parcels  $P_i$  and  $P_j$ . We are considering the set of bundles of all the ROIs in  $P_i$  connected with those in  $P_j$ . The correlation matrices are therefore the same size as the low-resolution connection matrix for all the data arising from any resolution.

These new matrices of each reprocessed subject data are correlated two by two to give five correlation coefficient matrices as well as their corresponding  $p$ -values. The cells in the matrices containing less than two values (corresponding to the set of numbers of bundles connecting the pair of the parcels) and



**Fig. 6.** Comparison of connection matrices. The red box plots show the correlation coefficients between connection matrices of same five subjects' scans processed twice to assess the method repeatability. The green ones are related to the same five subjects acquired twice to evaluate the intra-subjects variability. The blue box plots express the inter-subjects variability studied on 20 subjects. For the three tests, we observe a linear evolution of the correlation coefficients increasing from the highest resolutions to the lowest ones. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

those where the fibers are artificially removed with a mask<sup>1</sup> are not considered in the calculation on the correlation coefficient. Nevertheless, the cells where no bundles are found in both processings are completely correlated ( $cor = 1$ ) and considered. The average correlation between the five subjects double sets is then calculated considering only the locations where the correlation coefficient is significant for all the five double set (where the corresponding  $p$ -values are  $p < 5\%$  corrected for the 5 correlation matrices). These correlation matrices are illustrated in Fig. 9A. They show in both hemispheres very high correlations in most of connections in the matrices, which is consistent with the global correlation coefficients illustrated in the scatter plot of the reprocessed data.

(b) Same subject, scanned twice

Given that additional perturbation factors related to the MRI acquisition need to be taken into account, such as noise, head positioning and motion, susceptibility artifacts and head-coil sensitivity, the same subjects was scanned a second time, about one month later, following the same protocol. The two scans for every subject were processed independently.

The Pearson correlation coefficients between every pair of resulting connection matrices were calculated. The results presented in green in Fig. 6 show high correlation ranging from 0.976 to 0.874, though slightly lower than in the above section ( $p \ll 10e-5$ ). The same evaluations as presented in (a) are used in this section. Fig. 8 shows the resulting scatter plots as well as the Bland–Altman plot. As expected the repeatability measure is considerably affected by the scan noise. The straight lines at

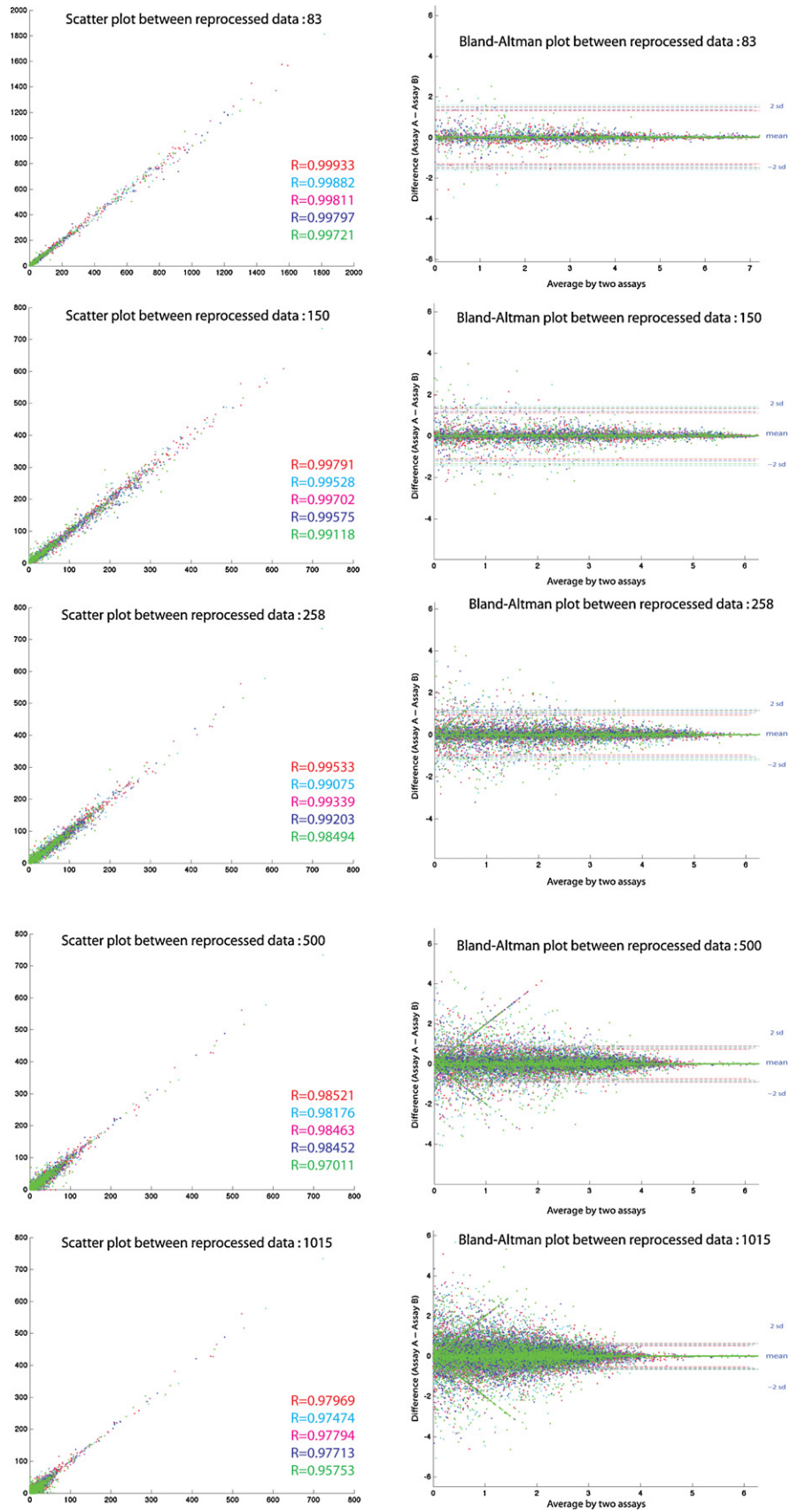
45° in the Bland–Altman plot, correspond to the matrix positions where one observation has a zero value (no connection) and the other has a connection present in the matrix. This is more visible in the rescanned data typically because of the diffusion noise. The same correlation coefficient matrices are calculated for the rescanned data double sets. This is performed as described above in the reprocessed data section, to highlight in this case the nodes where the correlation is significantly low in the same brain with different acquisitions. In Fig. 7B as noted before, only significant correlations for all subjects are shown. As expected, according to the correlations matrices the short bundles (in diagonal) are less robust than the long ones. This is potentially due to diffusion noise but also to the used tractography algorithm, and its corresponding parameters, such as the number of seed points or the number of considered direction of the ODF chosen here to perform an acceptable tractography.

For both experiments, the reprocessed and rescanned data, the correlation resulting from the two higher resolution connection matrices, seems to be better. It appears more dense than the ones given by the low-resolution data, where holes made the correlation matrices sparse. Nevertheless, this should be interpreted with caution. In fact, low-resolution matrices are more likely to contain parcels with only one or two ROIs, which are not considered in the calculation of the correlation as noted above.

### 3.2.2. Robustness over several subjects

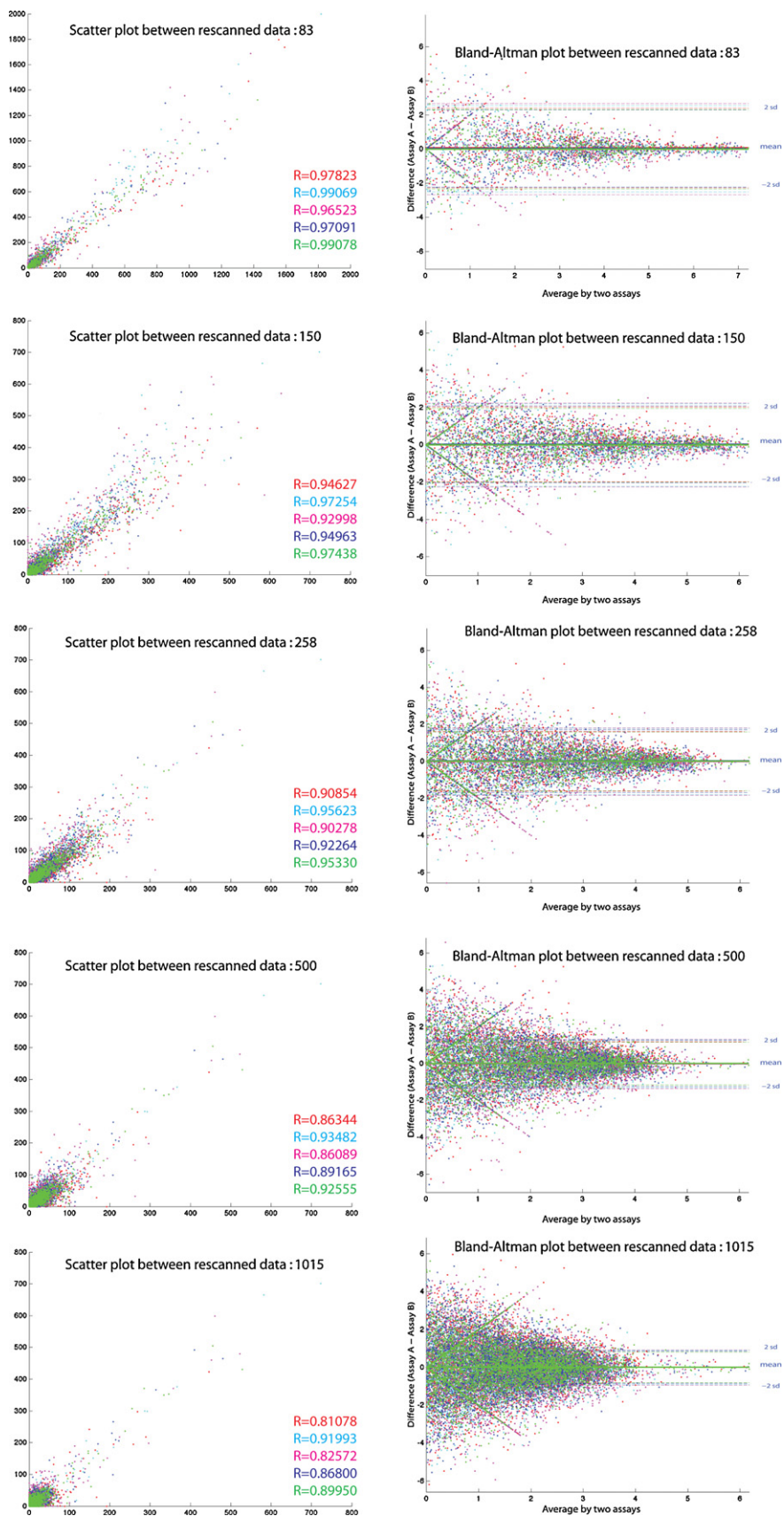
Quantitatively evaluating the quality of the cortical ROI matching over different subjects is a difficult task since cortical morphology varies strongly from one subject to the other, and therefore even experts have a hard time agreeing on the location of such small ROIs on the cortex. Therefore, we assessed the inter-variability between the 20 subjects by calculating the Pearson correlations between all pairs of their connection matrices on every

<sup>1</sup> The inter-hemispheric cells connecting all frontal part to pre-central, post-central and paracentral are forced to zeros; as well as the inter-hemispheric connections of the gray nuclei to the other ROIs. Only the homotopic connections of thalami, sub-thalami, hypo-campus and the amygdala are maintained.



**Fig. 7.** Same dataset processing repeatability assessment: (A) scatter plot of reprocessed data set of 5 subjects, and their correlation coefficient. Each color is corresponding to a subject dataset. (B) The Bland–Altman plot of the reprocessed datasets. The same color code than in (A) is adopted.





**Fig. 8.** Subject's repeatability assessment: (A) scatter plot of rescanned data set of 5 subjects, and their correlation coefficient. (B) The Bland–Altman plot of the rescanned datasets. The same color code than in (Fig. 7) is adopted.

scale. After Bonferroni correction for multiple testing, 50% of the correlation coefficient ranges from 0.724 to 0.767 for the highest resolution and between 0.939 and 0.958 for the lowest resolution ( $p \ll 10e-5$ ). These correlation coefficients presented in the blue box plot of Fig. 6 show a similarity between the brain connectivity among the group of subjects but also indicate an expected human inter-subject variability making the coefficients lower than those calculated in the previous sections for the same subjects. (For interpretation of the references to color in this text, the reader is referred to the web version of the article.)

To assess the local inter-subject variability of the method, the correlation matrices are calculated between the twenty subjects for each parcel and for each resolution. By averaging the pair-wise correlation matrices between each pair of twenty subjects and by keeping only significant correlations, we obtain for the four resolutions what we may call intervariability matrix. One interesting point is that the resolutions 250 and 125 seem to show the better results (Fig. 9C). In fact the highest ones contain many significantly low correlated connections, and the low-resolution one is sparse because the parcels do not contain enough regions to make the comparison possible.

The most important point to notice in this experiment is that the shape of the connectivity over the entire set of subjects is maintained. In fact even though the correlation coefficients vary, the high correlation values prove the stability of the connectivity between the twenty subjects.

It is important to note that for the inter-subject comparison and for the assessment of the variability, normalization for brain size is required. The value of each matrix cell is therefore divided by the number of voxels in the related white matter mask, corresponding to the number of seed points used for the fibers simulation.

### 3.3. ROIs registration reliability

In order to assess the ROIs registration accuracy, at all the scales, the following questions have been investigated: (1) is a particular cortical area always anatomically located at the same place over all the subjects? And do selected cortical landmarks on the different subjects fall in the same ROI labels? (2) if not, is the distance between the homologous regions important?

To investigate and answer these questions, several well-featured points scattered over the cortex were manually located by an expert (P.H.) on the template as well as on the 5 subject's cortices. The selection of these points was adhoc, chosen as visually identifiable by anatomical landmarks. These reference points on the template; are shown for both hemispheres in Fig. S1.

Let us call  $V_T$  the set of the 26 vertices selected by the expert on the template. These reference vertices were transferred on the 5 different subjects using the respective transformations already calculated for the registration between the template and the cortical surface of each subject. Let us denote by  $V_{Ti} i = \{1, \dots, 5\}$  the set of reference vertices after the transformation on the subject's cortices. The same expert has located the same cortical points on the 5 subjects and selected the corresponding vertices; each set of subject's points is called  $V_{SBji} i = \{1, \dots, 5\}$ .

If the label of the point  $i$  in  $V_{Ti}$  (reference vertices set after transformation on the subject's cortex) is the same label as the one of its homologous  $i$  in  $V_{SBji}$  for a given subject, we consider the matching level equal to one, otherwise to zero. For each resolution the matching level is calculated for all the points (26) projected on the 5 subject's cortices, the means and STD are, respectively  $0.65 \pm 0.12$ ,  $0.72 \pm 0.15$ ,  $0.74 \pm 0.1$ ,  $0.76 \pm 0.09$  and  $0.9 \pm 0.1$  for the resolution going from 1015 to 83.

Notice if the matching level seems low in the highest resolution, when the expected ROIs in  $V_{Ti}$  and its analogous in  $V_{SBji}$  are different, they are often neighbors.

To quantify this, the geodesic distance between every pairs of vertices in  $V_{Ti}$  and  $V_{SBji}$  has been calculated. The mean distance is  $5.24 \pm 2.42$  mm confirming the nearness of the ROIs, even if the matching is not always perfect.

## 4. Discussion

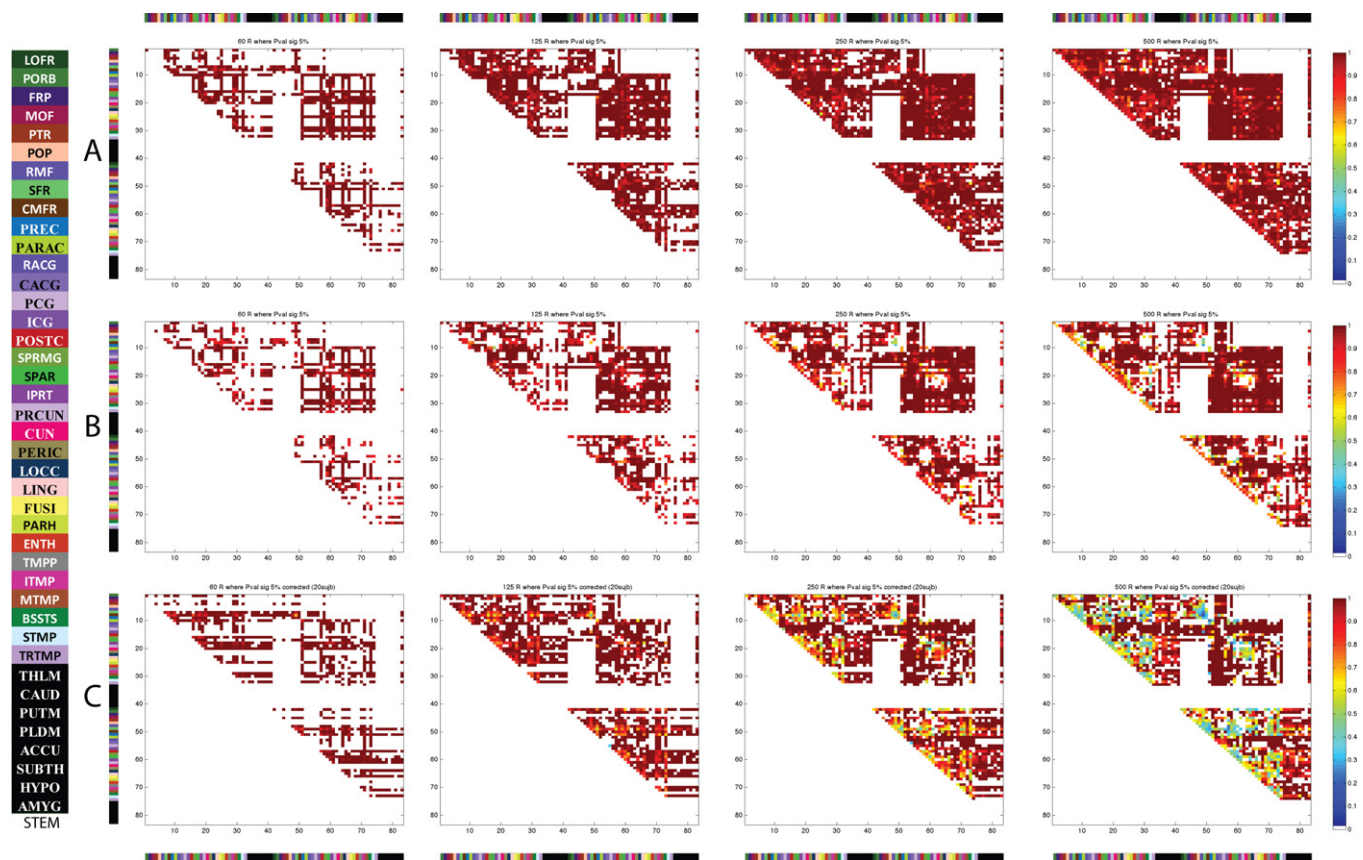
Over the last years it has become clear that MR based connectomic techniques are of the highest interest for the neuroscience community (Bassett et al., 2010; Bullmore and Sporns, 2009; Fornito et al., 2010; Gross, 2008; Hagmann et al., 2010; Zalesky et al., 2010), but methodological issues remains. The presented method is a contribution to tackle these issues. We showed step by step how to partition the cortex in a standard way such that ROIs are robustly placed on the same cortical surface across subjects, enabling the construction of whole brain normalized connection matrix at multiple scales, which can be averaged and compared over population of subjects.

Defining a connection matrix as presented here with a connection density measure, is only one way to characterize connectivity. Other tract properties can be computed to construct the matrix, such as the Fractional Anisotropy or the Mean Diffusion or even other values measured by MRI such as the T2 or the Magnetization Transfer Ratio (MTR) measured along the tracts, or even functional connectivity based on fMRI. The integration of such measurements will be worthwhile studying in the future.

The quantitative assessment of the performance of our methodology allows us to estimate what can and cannot be achieved. We saw that the level of correlation between two successive processing of same scans or of two scans of same subjects are around 0.9 which is reliable despite the level of complexity and the number of steps that are required to compute the connection matrix from the raw images. With this remark in mind we see that the correlations between two scans of the same subjects, but also overall 20 subjects for all resolutions, are indeed high, especially for resolution up to 133 or 241 ROIs. In fact for those latter resolutions, the ROIs registration was shown to be robust, since an expert selected the correct ROIs with success rates of 77% and 73%, respectively. Additional experiments show precisely the location in the matrix where the correlation coefficient is relatively high or low. They illustrate also the fact that short bundles are generally less robust than long bundles.

Depending on the study, going to smaller matrices might be beneficial because inter-subject correlation is highest (around 0.951). However, significant changes in connectivity may be lost due to large spatial averaging. On the other hand, using matrices that are too large may have, as a consequence, that variability due to registration errors and MRI noise is larger than the effective connectivity difference between two groups. Nevertheless, it has been shown that even the highest resolution connection matrices with the smallest ROIs have capacity to be at the source of powerful analyses (Fornito et al., 2010; Hagmann et al., 2008; Honey et al., 2009; Zalesky et al., 2010) (a) if one wants to study the connectivity in one single subject (b) if connectivity has to be characterized through local and global network measures, in which case averaging occurs through the network measure itself. In conclusion, the right scale depends on the particular question being investigated. This is the reason why a multiscale approach may be valuable.

In fact, the subdivision of the big parcels into smaller ROIs, normalized across subjects, enables to detail one given bundle and see its ramification and vice versa. The role of the node number is therefore important as it allows to give a multi-scale, embedded, description of the connectome. This paper is precisely about this,



**Fig. 9.** Intra and inter correlation coefficient matrices: (A) the average intra subject correlation matrix calculated for the five reprocessed subjects' double sets. (B) The average intra subject correlation matrix calculated for the five rescanned subjects' double sets. (C) The average inter subject correlation matrix for 20 subjects. Only significant ( $p < 0.05$ ) correlation coefficients are considered. Brown to yellow corresponds to decreasing correlation coefficient; and white corresponds to not applicable correlation coefficient. These matrices are calculated for resolution 1015, 500, 258 and 150. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

and about the robustness and reproducibility of such an embedded representation.

It is however, important to note that one of the limitations of this method is directly linked to the tractography tool adopted for the reconstruction of the virtual fibers. Depending on the chosen method this limitation is more or less sensitive to noise and partial volume effects. In the current work we used the simplest tractography method, which is the streamline algorithm adapted to work with multiple directions reconstructed with DSI. The main shortcoming of this method is that it exploits only a small portion (i.e., diffusion maxima of the ODFs) of the rich information provided by DSI (i.e., full 3D diffusion propagator). Concerning this, it is worth noting that the present method will automatically benefit from any future tractography method, which would fully exploit all the information available with DSI acquisitions. Because the goal of this paper is mainly to point out the potentials of the proposed method, this aspect has not been considered further. In the same context, tractography limitations affect the mapping of the human connectome: we may also capture aberrant connections arising from partial volume effects or noise. And, some connections running through large fiber crossing areas may not be mapped, especially small bundles crossing big well aligned bundles, like for instance the thalamo-frontal bundle generally deviated by the callosal bundles, in our experiments. However, this methodology provides us with a map of global connectivity which is representative of the major fiber pathways of the brain and will directly benefit from any advantage of other sophisticated tractography algorithms minimizing these limitations.

This normalized connection matrix opens up a whole range of opportunities for clinical studies; either for longitudinal healthy development analysis such as the development of connectivity with age, or for the investigation of connective disturbances in disease. Indeed, in many pathologies, such as schizophrenia or epilepsy, connectivity in some specific bundles is suspected of being affected while inflammatory processes may affect connectivity more globally. To investigate this kind of pathology the connection matrix provides a promising tool for performing a group versus group comparison if the difference in connectivity can be expected to be in the same brain areas, or throughout the brain. Please notice that especially, this multi-scale approach allows developing robust and powerful methods for statistical group comparison of connection matrices, that fully exploits this embedding of one resolution in the next one (see Meskaldji et al., 2011) that we published recently.

#### Acknowledgments

This work is partially supported by the Center for Biomedical Imaging (CIBM) of the Geneva-Lausanne Universities, the EPFL, and the foundations Leenaards and Louis-Jeantet, by an interdisciplinary grant of University of Lausanne and by a prospective researcher grant of Swiss National Science Foundation. O.S. was supported by the JS McDonnell Foundation. The authors would like to thank Prof. S. Morgenthaler for his precious collaboration as well as Dr. M. Saenz for English improvement.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jneumeth.2011.09.031.

## References

- Bassett DS, Brown JA, Deshpande V, Carlson JM, Grafton ST. Conserved and variable architecture of human white matter connectivity. *Neuroimage* 2010;54:1262–79.
- Behrens TE, Johansen-Berg H. Relating connective architecture to grey matter function using diffusion imaging. *Philos Trans R Soc Lond B Biol Sci* 2005;360:903–11.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–10.
- Broca P. Remarks on the seat of the faculty of articulated language following an observation of aphemia (loss of speech). *Bull Soc Anat* 1861;6:330–57.
- Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci* 2009;10:186–98.
- Clarke S, Riahi-Arya S, Tardif E, Eskenasy AC, Probst A. Thalamic projections of the fusiform gyrus in man. *Eur J Neurosci* 1999;11:1835–8.
- Conturo TE, Lori NF, Cull TS, Akbudak E, Snyder AZ, Shimony JS, et al. Tracking neuronal fiber pathways in the living human brain. *Proc Natl Acad Sci U S A* 1999;96:10422–7.
- Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 2006;31:968–80.
- Di Virgilio G, Clarke S, Pizzolato G, Schaffner T. Cortical regions contributing to the anterior commissure in man. *Exp Brain Res* 1999;124:1–7.
- Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1991;1:1–47.
- Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 2002;33:341–55.
- Fischl B, van der Kouwe A, Destrieux C, Halgren E, Segonne F, Salat DH, et al. Automatically parcellating the human cerebral cortex. *Cereb Cortex* 2004;14:11–22.
- Fornito A, Zalesky A, Bullmore ET. Network scaling effects in graph analytic studies of human resting-state fMRI data. *Front Syst Neurosci* 2010;4:22.
- Gall FJ, Spurzheim G. Anatomie et physiologie du système nerveux en général et du cerveau en particulier, avec des observations sur la possibilité de reconnaître plusieurs dispositions intellectuelles et morales de l'homme et des animaux par la configuration de leurs têtes. F. Schoell (Paris) ed; 1810–1819.
- Gong G, He Y, Concha L, Lebel C, Gross DW, Evans AC, et al. Mapping anatomical connectivity patterns of human cerebral cortex using in vivo diffusion tensor imaging tractography. *Cereb Cortex* 2009;19:524–36.
- Gong G, He Y, Concha L, Lebel C, Gross DW, Evans AC, et al. Mapping anatomical connectivity patterns of human cerebral cortex using in vivo diffusion tensor imaging tractography. *Cereb Cortex* 2008.
- Gross L. From structure to function: mapping the connection matrix of the human brain. *PLoS Biol* 2008;6:e164, doi:10.1371.
- Hagmann P. From diffusion MRI to brain connectomics. Lausanne: Ecole Polytechnique Fédérale de Lausanne (EPFL); 2005.
- Hagmann P, Cammoun L, Gigandet X, Gerhard S, Grant E, Wedeen V, Meuli R, Thiran JP, Honey CJ, Sporns O. MR connectomics: principles and challenges. *J Neurosci Methods* 2010;194(1):34–45.
- Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, et al. Mapping the structural core of human cerebral cortex. *PLoS Biol* 2008;6:e159.
- Hagmann P, Kurlant M, Gigandet X, Thiran JP, Wedeen VJ, Meuli R, et al. Mapping human whole-brain structural networks with diffusion MRI. *PLoS One* 2007;2:e597.
- Hagmann P, Reese TG, Tseng W-YI, Meuli R, Thiran JP, Wedeen VJ. Diffusion spectrum imaging tractography in complex cerebral white matter: an investigation of the centrum semiovale. In: International society for magnetic resonance in medicine, ISMRM twelfth scientific meeting; 2004. p. 623.
- Hagmann P, Thiran JP, Jonasson L, Vandergheynst P, Clarke S, Maeder P, et al. DTI mapping of human brain connectivity: statistical fibre tracking and virtual dissection. *Neuroimage* 2003;19:545–54.
- Hilgetag CC, Kaiser M. Clustered organization of cortical connectivity. *Neuroinformatics* 2004;2:353–60.
- Hilgetag CC, O'Neill MA, Young MP. Hierarchical organization of macaque and cat cortical sensory systems explored with a novel network processor. *Philos Trans R Soc Lond B Biol Sci* 2000;355:71–89.
- Honey CJ, Sporns O, Cammoun L, Gigandet X, Thiran JP, Meuli R, et al. Predicting human resting-state functional connectivity from structural connectivity. *Proc Natl Acad Sci U S A* 2009;106:2035–40.
- Iturria-Medina Y, Canales-Rodriguez EJ, Melie-Garcia L, Valdes-Hernandez PA, Martinez-Montes E, Aleman-Gomez Y, et al. Characterizing brain anatomical connections using diffusion weighted MRI and graph theory. *Neuroimage* 2007;36:645–60.
- Iturria-Medina Y, Sotero RC, Canales-Rodriguez EJ, Aleman-Gomez Y, Melie-Garcia L. Studying the human brain anatomical network via diffusion-weighted MRI and Graph Theory. *Neuroimage* 2008;40:1064–76.
- Klein JC, Behrens TE, Robson MD, Mackay CE, Higham DJ, Johansen-Berg H. Connectivity-based parcellation of human cortex using diffusion MRI: establishing reproducibility, validity and observer independence in BA 44/45 and SMA/pre-SMA. *Neuroimage* 2007;34:204–11.
- Li Y, Liu Y, Li J, Qin W, Li K, Yu C, et al. Brain anatomical network and intelligence. *PLoS Comput Biol* 2009;5:e1000395.
- Meskaldji DE, Ottet MC, Cammoun L, Hagmann P, Meuli R, Eliez S, et al. Adaptive strategy for the statistical analysis of connectomes. *PLoS One* 2011;6:e23009.
- Mori S, Crain BJ, Chacko VP, van Zijl PC. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Ann Neurol* 1999;45:265–9.
- Raichle ME, Snyder AZ. A default mode of brain function: a brief history of an evolving idea. *Neuroimage* 2007;37:1083–90, discussion 97–99.
- Schmahmann JD, Pandya DN. *Fiber pathways of the brain*; 2006, New York.
- Schmahmann JD, Pandya DN, Wang R, Dai G, D'Arceuil HE, de Crespigny AJ, et al. Association fibre pathways of the brain: parallel observations from diffusion spectrum imaging and autoradiography. *Brain* 2007;130:630–53.
- Sporns O. <http://www.scholarpedia.org/article/Connectome> 2008.
- Sporns O, Tononi G, Kötter R. The human connectome: a structural description of the human brain. *PLoS Comput Biol* 2005;1:e42.
- Sporns O, Zwi JD. The small world of the cerebral cortex. *Neuroinformatics* 2004;2:145–62.
- Stephan KE, Riera JJ, Deco G, Horwitz B. The brain connectivity workshops: moving the frontiers of computational systems neuroscience. *Neuroimage* 2008;42:1–9.
- Thottakara P, Lazar M, Johnson SC, Alexander AL. Application of Brodmann's area templates for ROI selection in white matter tractography studies. *Neuroimage* 2006;29:868–78.
- Tomassini V, Jbabdi S, Klein JC, Behrens TE, Pozzilli C, Matthews PM, et al. Diffusion-weighted imaging tractography-based parcellation of the human lateral premotor cortex identifies dorsal and ventral subregions with anatomical and functional specializations. *J Neurosci* 2007;27:10259–69.
- Van Leemput K, Maes F, Vandermeulen D, Suetens P. Automated model-based tissue classification of MR images of the brain. *IEEE Trans Med Imaging* 1999;18:897–908.
- Wedeen V. Diffusion anisotropy white matter tracts. In: *Second int. conf. funct. mapp. Human brain*; 1996.
- Wedeen VJ, Hagmann P, Tseng WY, Reese TG, Weisskoff RM. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magn Reson Med* 2005;54:1377–86.
- Wedeen VJ, Wang RP, Schmahmann JD, Benner T, Tseng WY, Dai G, et al. Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. *Neuroimage* 2008;41:1267–77.
- Wernicke C. *Grundriss der Psychiatrie*. Leipzig: Verlag; 1906.
- Zaidel E, Aboitiz F, Clarke J. Sexual dimorphism in interhemispheric relations: anatomical-behavioral convergence. *Biol Res* 1995;28:27–43.
- Zalesky A, Fornito A, Harding IH, Cocchi L, Yücel M, Pantelis C, et al. Whole-brain anatomical networks: does the choice of nodes matter? *Neuroimage* 2010;50:970–83.