

Multi-way Regression Reveals Backbone of Macaque Structural Brain Connectivity in Longitudinal Datasets

Tuo Zhang¹(✉), Xiao Li¹, Lin Zhao¹, Xintao Hu¹, Tianming Liu²,
and Lei Guo¹

¹ Brain Decoding Research Center,
Northwestern Polytechnical University, Xi'an, Shaanxi, China
tuozhang@nwpu.edu.cn

² Computer Science Department, The University of Georgia, Athens, GA, USA

Abstract. Brain development has been an intriguing window to study the dynamic formation of brain features at a variety of scales, such as cortical convolution and white matter wiring diagram. However, recent studies only focused on a few specific fasciculus or several region. Very few studies focused on the development of macro-scale wiring diagrams due to the lack of longitudinal datasets and associated methods. In this work, we take the advantage of recently released longitudinal macaque MRI and DTI datasets and develop a novel multi-way regression method to model such datasets. By this method, we extract the backbone of structural connectome of macaque brains and study the trajectory of its development over time. Graphic statistics of these backbone connectomes demonstrate their soundness. Our findings are consistent with the reports in the literatures, suggesting the effectiveness and promise of this framework.

Keywords: Macaque brain connectivity · Longitudinal study · Multiway regression

1 Introduction

Nowadays, connectome has been widely studied in basic brain research and clinical studies due to its importance in understanding structural and functional brain architectures from a global perspective [1, 2]. Studying the development of such connectomes could play a critical role in understanding a variety of time-dependent issues, such as the mechanisms of cortical convolution, the emergence of cognition, the possible onset of brain abnormality, and etc. [3]. However, most previous connectome studies are limited to adult brains because of the unavailability of longitudinal datasets [1, 2].

Recently, a longitudinal UNC-Wisconsin neurodevelopment rhesus MRI database is released (http://www.nitrc.org/projects/uncuw_macdevmri/), which includes T1-weighted MRI and diffusion tensor imaging (DTI) data, providing a unique opportunity for studying how brain structures of newborns develop into their infancy, juvenile and adulthood. Studies based on such longitudinal datasets could significantly complement the previous structural brain development studies in three ways: (1) many previous

imaging data studies were voxel based. Voxel-wise attributes such as white matter volume growth rates [4, 5] were well studied but they provided limited clues to the development of axonal pathways; (2) axonal pathway related studies were only limited to a certain fasciculus of interest, such as the developments in Area 17&18 connections [6]. It is difficult to make up a big picture of the evolvement of an axonal network architecture, or a connectome, from these scattered profiles; (3) although dissection studies based on data modalities such as microscopic histology data and meso-scale tract-tracing data were widely available [7], it is not possible at this moment to collect longitudinal data or reconstruct a global scale brain connectome via these techniques.

Therefore, we take the advantages of the UNC-Wisconsin longitudinal macaque dataset and attempt to study the postnatal development of global wiring diagram. We construct longitudinal connective matrices for a subject at four different time points. A novel multi-way regression method is proposed to extract the ‘backbone’ connective network by using the matrix of the first time point as the baseline. This process is repeated for the other time points, such that a trajectory of ‘backbone’ connective networks are obtained. We use a variety of graphic statistics, such as node-wise strength and betweenness and efficiency of networks, to demonstrate the soundness of these ‘backbone’ networks. The effectiveness of the method also suggests its potential in future studies of brain functions and abnormality.

2 Materials and Methods

2.1 Overview

Generally, we use T1-weighted MRI, DTI data and a brain map to construct a global cortical connectome, a connective matrix, for a subject at a time point (the first three columns in Fig. 1). All subjects are divided to groups according to the time points of scans. These matrices are stretched to connective feature vectors, and these vectors within one time point compose a feature matrix (the 4th column). The feature matrix of the t_1 is taken as the baseline, to which the other feature matrices are group-wisely regressed to produce weight vectors \mathbf{v}_s via the proposed multi-way regression method. The weight vectors can be transformed back to connective matrices (the right-most row). Because of the sparse term (see Eq. (6) in Sect. 2.3), the nonzero values in these weight matrices could imply that they are ‘preserved’ connections over time, and thus taken as ‘backbone’ connections associated to t_j . Notably, we will not differentiate the weight vectors and their corresponding matrices format in the following sections. It is also worth noting that Fig. 1 only illustrates the pipeline for t_j . The method is progressively applied to the other time points in a similar way, such that a trajectory of such ‘backbone’ connective maps are obtained.

2.2 Data Description and Preprocessing

T1-weighted MRI and DTI data in the UNC-Wisconsin neurodevelopment rhesus MRI database are used. Four time intervals are defined as four time points: 1–4 months, 5–8 months, 9–12 months and 13–16 months. Ten subjects that have data of all the four

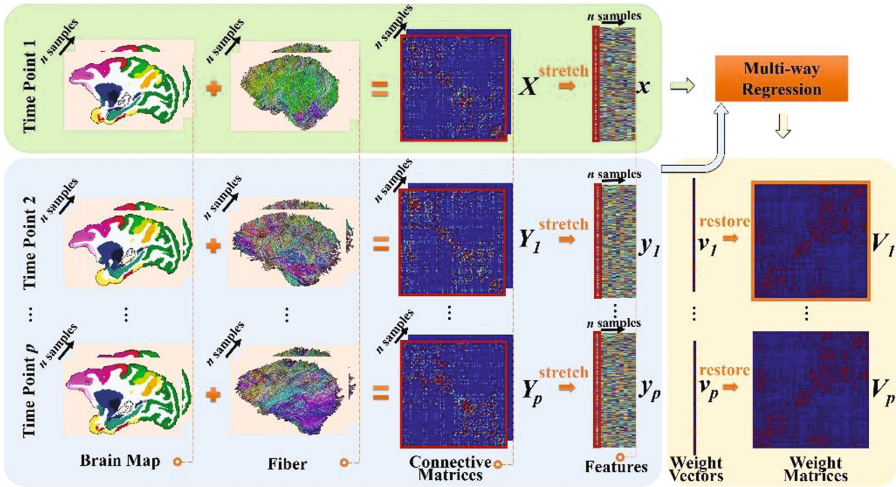


Fig. 1. Flow chart of the analysis pipeline in this paper. Data of the 1st time point (t_l for short), for example, in the green frame is used as baseline, to which data of other time points in the blue frame is regressed simultaneously *via* the multi-way regression method detailed in Sect. 2.3. The yellow frame highlights the results. Please refer to main text for detailed explanations.

time points are selected in this work. The basic parameters for diffusion data acquisition are: resolution of $0.65 \times 0.65 \times 1.3 \text{ mm}^3$, a matrix of $256 \times 256 \times 58$, diffusion-weighting gradients applied in 120 directions and b value of 1000 s/mm^2 , ten images without diffusion weighting. For each DTI scan, the corresponding T1 weighted MRI data in the dataset has been registered to the UNC Primate Brain Atlas space [8]. The resolution of this atlas space is $0.27 \times 0.27 \times 0.27 \text{ mm}^3$ and a matrix of $300 \times 350 \times 250$.

DTI data is defined as intra-subject standard space. T1 weighted MRI data is nonlinearly warped to the FA map via FSL-fnirt [9]. Two axonal pathway orientations for each voxel in DTI data is estimated via BedpostX in FSL 5, because it was suggested that b -values of 4000 would be required to resolve a 3 fiber orthogonal system [10]. Deterministic white matter fiber streamlines (1×10^4 fiber tracts for each DTI data) are reconstructed via DSStudio [11] thereafter. Because the selection of fiber tracking parameters is not the focus of this paper, FA and angular threshold are empirically determined as 0.1 and 60° in that small FA value for primate brain was suggested [12].

To construct structural connective matrices from DTI data, we adopt the paxinos rhesus monkey atlas [13] as the brain map, the brain sites of which are used as nodes. Because this brain map is present in T1-weighted MRI atlas space, we firstly use FSL-fnirt [9] to register the atlas T1-weighted MRI data to the one of an individual, and further register it to the DTI space (FA map) of the individual. The brain map is transposed to the individual DTI space, accordingly. Cortico-cortical connections are of our major interest here. Thus, 152 cortical brain sites on each hemisphere are used and both ipsilateral and contralateral connections are considered. The number of fiber tracts

that connect two brain sites are defined as the connective strength. Examples of such connective matrices, \mathbf{X} and \mathbf{Y}_s , are shown in Fig. 1. The connective strength within each individual is normalized by being divided by the total fiber number. We stretch the upper triangle of a connective matrix into a vector, and define it as a connective feature. A pile of such features of ten subjects within one time point gives a feature matrix (\mathbf{x} and \mathbf{y}_s in Fig. 1).

2.3 Multi-way Regression

We have a collection of feature matrices ($\mathbf{x}, \mathbf{y}_1, \dots, \mathbf{y}_p$) where p is the number of time points. \mathbf{x} and \mathbf{y}_k are feature matrices, consisting of (x^1, \dots, x^n) and (y_k^1, \dots, y_k^n) , where n is the number of subjects and x^k is the feature vector converted from connective matrix (see Fig. (1)). Our objective is to use \mathbf{x} as a baseline and find a group of weights $(\mathbf{v}_1, \dots, \mathbf{v}_p)$, to which $(\mathbf{y}_1, \dots, \mathbf{y}_p)$ are projected (Eq. (1)), respectively, such that new \mathbf{y}_k s shown below are group-wisely similar to \mathbf{x} . We will use \mathbf{y} in the following equations to denote those before the projection.

$$\mathbf{y}_k^{new} \rightarrow \langle \mathbf{v}_1, \mathbf{y}_k^{old} \rangle \quad (1)$$

The similarity is defined as correlation. Therefore, the function to be maximized is

$$\max_{\mathbf{v}_1, \dots, \mathbf{v}_p} \sum_{k=1}^p \text{corr}(\mathbf{x}, \mathbf{y}_k) + \sum_{k=1}^p \sum_{l=k+1, l \neq k}^p \text{corr}(\mathbf{v}_k, \mathbf{v}_l) \quad (2)$$

where the second term is to maximize the similarity among \mathbf{v}_k . With this term, \mathbf{y} s will be regressed to \mathbf{x} in a group-wise manner. We define the correlation as

$$\text{corr}(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x}, \mathbf{y} \rangle / \|\mathbf{x}\| \|\mathbf{y}\| \quad (3)$$

and covariance matrix of (\mathbf{x}, \mathbf{y}) as

$$\mathbf{C}(\mathbf{x}, \mathbf{y}) = \begin{bmatrix} C_{xx} & C_{xy} \\ C_{yx} & C_{yy} \end{bmatrix} \quad (4)$$

where C_{xx} and C_{yy} are within-set covariance matrices and $C_{xy} = C_{yx}'$ are between-sets covariance matrices, we can rewrite Eq. (2) as

$$\max_{\mathbf{v}_1, \dots, \mathbf{v}_p} \sum_{k=1}^p \frac{\mathbf{u}' C_{xy} \mathbf{v}_k}{\sqrt{\mathbf{u}' C_{xx} \mathbf{u} \mathbf{v}_k' C_{yy} \mathbf{v}_k}} + \sum_{k=1}^p \sum_{l=k+1, l \neq k}^p \frac{\mathbf{v}_l' \mathbf{v}_k}{\sqrt{\mathbf{v}_l' \mathbf{v}_l \mathbf{v}_k' \mathbf{v}_k}} \quad (5)$$

where $\mathbf{u} = \mathbf{1}$, because \mathbf{x} is used as a baseline. The corresponding Lagrangian is

$$L(\mathbf{v}_1, \dots, \mathbf{v}_p) = \sum_{k=1}^p \mathbf{u}' C_{xy_k} \mathbf{v}_k - \frac{p}{2} \mathbf{u}' C_{xx} \mathbf{u} - \frac{1}{2} \sum_{k=1}^p \mathbf{v}_k' C_{y_k y_k} \mathbf{v}_k + \sum_{k=1}^p \sum_{l=k+1, l \neq k}^p \mathbf{v}_l' \mathbf{v}_k - \frac{(p-1)}{2} \sum_{k=1}^p \mathbf{v}_k' \mathbf{v}_k - \lambda \sum_{k=1}^p \|\mathbf{v}_k\|_1 \tag{6}$$

Equation (6) is our objective function. Because we have long and sparse feature vectors \mathbf{x} and \mathbf{y} , ℓ_1 norm, $\lambda \sum_{k=1}^p \|\mathbf{v}_k\|_1$, is added to Eq. (6) to give sparse \mathbf{V}_{k^s}

By computing derivatives in respect to \mathbf{v}_k , we have

$$\frac{\partial L}{\partial \mathbf{v}_k} = C_{y_k x} \mathbf{u} - p C_{y_k y_k} \mathbf{v}_k + \sum_{k=1}^p \sum_{l=k+1, l \neq k}^p \mathbf{E} \mathbf{v}_l - (p-1) \mathbf{E} \mathbf{v}_k = 0 \tag{7}$$

$$(p(C_{y_k y_k} + \mathbf{E}) - \mathbf{E}) \mathbf{v}_k = C_{y_k x} \mathbf{u} + \sum_{k=1}^p \sum_{l=k+1, l \neq k}^p \mathbf{E} \mathbf{v}_l \tag{8}$$

Equation (8) has the linear form of $\mathbf{A} \mathbf{x} = \mathbf{b}$, and is used to iteratively update \mathbf{v}_k s. The algorithm is summarized and executed as follows:

```

Initialize  $\mathbf{v}_{k^s}$ 
While  $|\mathbf{L}(i) - \mathbf{L}(i-1)| > \text{stop\_threshold}$  &  $\text{iter} > \text{max\_iteration}$ 
  For  $k = 1, k \leq p, k = k+1$ 
    update  $\mathbf{v}_k$  by Eq.(8)
  End For
  update  $\mathbf{L}(i)$  by Eq.(6).
   $i = i + 1$ 
End While

```

We convert weight vectors \mathbf{v}_k to matrix format \mathbf{V}_{k^s} (see Fig. 1). The values of these matrices element represent their contribution to the similarity between \mathbf{Y}_k and \mathbf{X} . Because of the sparse term introduced, nonzero elements in \mathbf{V}_k indicate the corresponding connections in \mathbf{Y}_k are preserved from \mathbf{X} . It is worth noting that \mathbf{V}_{k^s} could be similar to each other due to the constraint $\text{corr}(\mathbf{v}_k, \mathbf{v}_l)$ in the objective function Eq. (2). Therefore, only the \mathbf{V}_1 corresponding to \mathbf{y}_1 is used as a representative (highlighted by an orange frame in Fig. 1) to interpret the results, defined as \mathbf{V}_{t_1} . Nonzero elements in \mathbf{V}_{t_1} could imply that the corresponding connections are preserved over all time points.

We execute this method in a progressive manner. For the first run, \mathbf{x} is the feature matrix of t_1 . $(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3)$ correspond to those from t_2 to t_3 . Unit vectors are used to initialize \mathbf{V}_{k^s} for this run. For the second run, \mathbf{x} is the feature matrix of t_2 . $(\mathbf{y}_1, \mathbf{y}_2)$ correspond to those from t_3 to t_4 . To initialize \mathbf{V}_{k^s} , we replace an element of unit vectors with 0 if its counterpart in \mathbf{V}_{t_1} from the previous run is nonzero, such that connections not preserved (or new to t_1) are considered with higher priority. Following this scheme, we obtain a series of representative $(\mathbf{V}_{t_1}, \mathbf{V}_{t_2}, \mathbf{V}_{t_3})$ for $t_1 \sim t_3$.

2.4 Graphic Statistics

In order to further explore the property and evaluate the validity of the extracted connective maps V_S , we adopt a variety of graphic statistics at different scales, such as node-wise measurements (strength and betweenness) and connectome-wise one (network efficacy). For node-wise measurements, we define an overlap ratio to compare two connective matrices. By taking node strength for example, we firstly compute the strength for each node of matrices X and Y , and obtain two strength vectors s_X and s_Y . They are then sorted in a descending order, denoted by s'_X and s'_Y . By taking X as a baseline matrix, we generate an overlap curve, the k^{th} value of which is defined as the length of intersection of the first k elements of s'_X and s'_Y divided by the length of s'_X . The overlap ratio is defined as the area under this overlap curve.

3 Results

3.1 Effectiveness of the Multi-way Regression Method

The λ in the objective function (Eq. (6)) regulates the sparsity of weight vectors V_{ks} and is optimized via a five-fold cross-validation scheme such that the objective function is maximized. To demonstrate the stability of the method on the used dataset, we divide the ten subjects into two groups (five subjects in each one with no overlap) and separately execute the algorithms on them. 250 permutations of such divisions are used. On average, the similarity (Pearson correlation) of the resulting weight vectors between two groups is 0.70 ± 0.34 , demonstrating the robustness of the methods on this dataset. It is worth noting that the resulting weight vectors on five subjects are definitely different from those on all ten subjects. This robustness experiment implies that our method can produce reproducible results on ten subjects. Therefore, the following analyses are all based on ten subjects.

3.2 Development of Connective Connectome of Infant Macaque Brains

Figure 2(a) shows the weight matrices on the left-most column. They are obtained by using connective matrices of t_1 , t_2 and t_3 as baselines, respectively. Connections exclusive to a certain time point are highlighted with different colors by comparisons among them. Backbone connections are represented by nonzero elements in them. They are thus used as masks to screen the backbone connections from the original connective matrices other than the baseline ones. For example, V_{t_1} associated to t_1 is used to mask the original connective matrices from t_2 to t_4 (the second row). We use the Pearson correlation coefficients to measure the similarity among the original connective matrices. They are compared to those obtained from the masked counterparts. Taking the second row for example, the Pearson correlation coefficients are computed between the masked matrices and the original t_1 matrix in the first row (Table 1). These results suggest that the similarity between the masked connective matrices and the baseline ones are significantly raised regarding to small p -values via t -test, suggesting the soundness of the extracted backbone connectomes. The soundness of these backbone

connectomes are also demonstrated by reports in the literatures. For example, cyan connections which are associated to and exclusive to t_3 strongly originate or terminate in frontal/temporal/parietal cortices (Fig. 2(b)) but not in occipital cortex. These results are consistent with previous reports that occipital cortex and white matters associated with it stop growing at this developmental stage (9 months–12 months) while the other cortices and the white matters associated with them have not reached their growth rate peaks yet [4, 5].

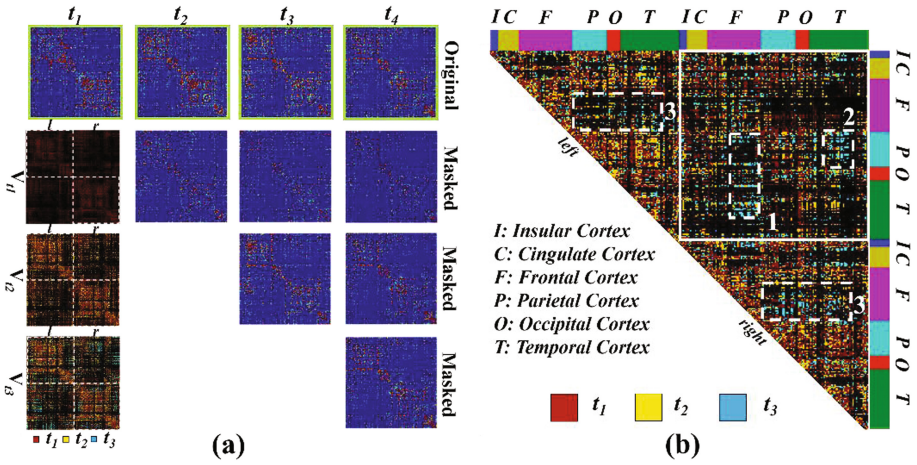


Fig. 2. (a) The three weight matrices V_S associated to $t_1 \sim t_3$ are on the left column. Elements exclusive to a certain time point are highlighted by different colors. Original connective matrices of four time points from one example subject are compared to those masked by the weight matrices in the left-most column. Interpretations are referred to the text. (b) Columns and rows of V_{β_3} in (a) are sorted in lobe order. White solid box separates ipsilateral and contralateral connections. White dashed boxes highlight groups of connections exclusive to t_3 .

To further demonstrate the validity of those backbone connectomes, a variety of graphic statistics in Sect. 2.4 are measured. We compute node strength/betweenness and network efficiency for all connective matrices in Fig. 2. The node-wise measurement overlap ratios defined in Sect. 2.4 between the masked connective matrices and their original counterparts as well as the global efficiency ratios are reported in Table 2. The network efficiency values obtained from the masked matrices are normalized by being divided by those obtained from their original counterparts. Taking the 2nd row in Table 2 (the 2nd row in Fig. 2) for example, we find that all measurements of the masked connective matrices from t_2 to t_4 remain at the same level, suggesting that the backbone connectomes are stable over time. This observation applies to the other rows as well, though these values increase after more connections are incorporated. Altogether, these results suggest that the development of connectome might prefer progressively making changes on the originally blank locations to strengthening the developed connections, which is one of the major neuroscientific insights achieved by this study.

Table 1. Evaluate the soundness of backbone connectomes by means of comparison of Pearson correlation coefficients between connective matrices. *P*-values of *t*-test are reported which are conducted on the corresponding cells on the two panels. O: original; M: masked.

	O2	O3	O4	M1 (<i>p</i> -value)	M2 (<i>p</i> -value)	M3 (<i>p</i> -value)
O1	0.27 ± 0.11	0.31 ± 0.13	0.28 ± 0.15	0.35 ± 0.13 2.63×10^{-4}	0.41 ± 0.16 1.72×10^{-5}	0.37 ± 0.17 3.00×10^{-5}
O2	–	0.41 ± 0.13	0.36 ± 0.12	–	0.44 ± 0.14 7.92×10^{-4}	0.39 ± 0.13 3.89×10^{-5}
O3	–	–	0.48 ± 0.10	–	–	0.50 ± 0.11 1.16×10^{-4}

Table 2. Strength & betweenness overlap ratios and network efficiency ratios of the masked connective matrices.

	Strength				Betweenness				Network efficiency			
	<i>t</i> ₁	<i>t</i> ₂	<i>t</i> ₃	<i>t</i> ₄	<i>t</i> ₁	<i>t</i> ₂	<i>t</i> ₃	<i>t</i> ₄	<i>t</i> ₁	<i>t</i> ₂	<i>t</i> ₃	<i>t</i> ₄
M1		0.92	0.92	0.91		0.73	0.75	0.72		0.74	0.75	0.73
M2			0.96	0.96			0.86	0.81			0.94	0.95
M3				0.97				0.87				0.97

4 Discussions and Conclusions

In this work, we proposed a novel multi-way regression method to extract developmental backbone connectomes on the longitudinal MRI and DTI macaque datasets. The effectiveness of those backbone connectomes is demonstrated by reproducibility studies, graphic statistics and previous reports. Studies on the trajectory of backbone connectomes could help reveal a possible structural wiring development mechanism, which would be further validated by other fine-grained data modalities and applied to abnormal brain datasets in the future. The objective of this study is in line with other developmental –omics studies, such as longitudinal functional connectome and genomics datasets recently released by the Allen Institute for Brain Science. An integrative analysis on those longitudinal –omics studies is also of our interest in the future.

References

1. Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Uğurbil, K.: WU-Minn HCP consortium.: the WU-Minn human connectome project: an overview. *Neuroimage* **80**, 62–79 (2013)
2. Jiang, T.: Brainnetome: a new -ome to understand the brain and its disorders. *Neuroimage* **80**, 263–272 (2013)
3. Innocenti, G.M., Price, D.J.: Exuberance in the development of cortical networks. *Nat. Rev. Neurosci.* **6**(12), 955–965 (2005)
4. Malkova, L., Heuer, E., Saunders, R.C.: Longitudinal magnetic resonance imaging study of rhesus monkey brain development. *Eur. J. Neurosci.* **24**(11), 3204–3212 (2006)

5. Liu, C., Tian, X., Liu, H., Mo, Y., Bai, F., Zhao, X., Ma, Y., Wang, J.: Rhesus monkey brain development during late infancy and the effect of phencyclidine: a longitudinal MRI and DTI study. *Neuroimage* **107**, 65–75 (2015)
6. Kennedy, H., Bullier, J., Dehay, C.: Transient projection from the superior temporal sulcus to area 17 in the newborn macaque monkey. *Proc. Natl. Acad. Sci. USA*. **86**(20), 8093–8097 (1989)
7. Webster, M.J., Bachevalier, J., Ungerleider, L.G.: Transient subcortical connections of inferior temporal areas TE and TEO in infant macaque monkeys. *J. Comp. Neurol.* **352**(2), 213–226 (1995)
8. Styner, M., Knickmeyer, R., Joshi, S., Coe, C., Short, S.J., Gilmore, J.: Automatic brain segmentation in rhesus monkeys. In: *Proceedings of SPIE Medical Imaging*, vol. 6512 (2007). 65122 L1-8
9. Andersson, J.L.R., Jenkinson, M., Smith, S.: Non-linear registration, aka spatial normalisation. FMRIB Technical report TR07JA2 (2010)
10. Behrens, T.E., Berg, H.J., Jbabdi, S., Rushworth, M.F., Woolrich, M.W.: Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage* **34** (1), 144–155 (2007)
11. Basser, P.J., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A.: In vivo fiber tractography using DT-MRI data. *Magn. Reson. Med.* **44**, 625–632 (2000)
12. Dauguet, J., Peled, S., Berezovskii, V., Delzescaux, T., Warfield, S.K., Born, R., Westin, C. F.: Comparison of fiber tracts derived from in vivo DTI tractography with 3D histological neural tract tracer reconstruction on a macaque brain. *Neuroimage* **37**, 530–538 (2007)
13. Paxinos, G., Huang, X., Toga, A.W.: *The Rhesus Monkey Brain in Stereo-taxic Coordinates*. Academic Press, San Diego (1999)