

Discovering Cortical Folding Patterns in Neonatal Cortical Surfaces Using Large-Scale Dataset

Yu Meng^{1,2}, Gang Li², Li Wang², Weili Lin², John H. Gilmore³,
and Dinggang Shen^{2(✉)}

¹ Department of Computer Science, University of North Carolina at Chapel Hill,
Chapel Hill, NC, USA

² Department of Radiology and BRIC, University of North Carolina
at Chapel Hill, Chapel Hill, NC, USA
dinggang_shen@med.unc.edu

³ Department of Psychiatry, University of North Carolina at Chapel Hill,
Chapel Hill, NC, USA

Abstract. The cortical folding of the human brain is highly complex and variable across individuals. Mining the major patterns of cortical folding from modern large-scale neuroimaging datasets is of great importance in advancing techniques for neuroimaging analysis and understanding the inter-individual variations of cortical folding and its relationship with cognitive function and disorders. As the primary cortical folding is genetically influenced and has been established at term birth, neonates with the minimal exposure to the complicated postnatal environmental influence are the ideal candidates for understanding the major patterns of cortical folding. In this paper, for the first time, we propose a novel method for discovering the major patterns of cortical folding in a large-scale dataset of neonatal brain MR images ($N = 677$). In our method, first, cortical folding is characterized by the distribution of sulcal pits, which are the locally deepest points in cortical sulci. Because deep sulcal pits are genetically related, relatively consistent across individuals, and also stable during brain development, they are well suitable for representing and characterizing cortical folding. Then, the similarities between sulcal pit distributions of any two subjects are measured from spatial, geometrical, and topological points of view. Next, these different measurements are adaptively fused together using a similarity network fusion technique, to preserve their common information and also catch their complementary information. Finally, leveraging the fused similarity measurements, a hierarchical affinity propagation algorithm is used to group similar sulcal folding patterns together. The proposed method has been applied to 677 neonatal brains (the largest neonatal dataset to our knowledge) in the central sulcus, superior temporal sulcus, and cingulate sulcus, and revealed multiple distinct and meaningful folding patterns in each region.

1 Introduction

The human cerebral cortex is a highly convoluted and complex structure. Its cortical folding is quite variable across individuals (Fig. 1). However, certain common folding patterns exist in some specific cortical regions as shown in the classic textbook [1], which examined 25 autopsy specimen adult brains. Mining the major representative patterns of cortical folding from modern large-scale datasets is of great importance in advancing techniques for neuroimaging analysis and understanding the inter-individual variations of cortical folding and their relationship with structural connectivity, cognitive function, and brain disorders. For example, in cortical surface registration [2], typically a single cortical atlas is constructed for a group of brains. Such an atlas may not be able to reflect some important patterns of cortical folding, due to the averaging effect, thus leading to poor registration accuracy for some subjects that cannot be well characterized by the folding patterns in the atlas. Building multiple atlases, with each representing one major pattern of cortical folding, will lead to boosted accuracy in cortical surface registration and subsequent group-level analysis.

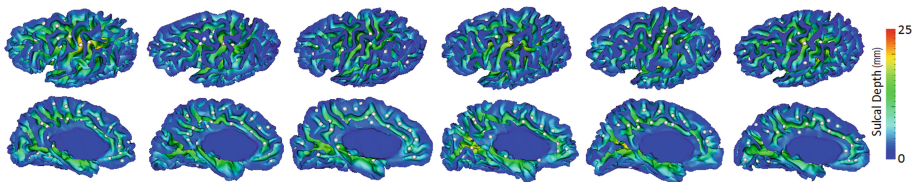


Fig. 1. Huge inter-individual variability of sulcal folding patterns in neonatal cortical surfaces, colored by the sulcal depth. Sulcal pits are shown by white spheres.

To investigate the patterns of cortical folding, a clustering approach has been proposed [3]. This approach used 3D moment invariants to represent each sulcus and used the agglomerative clustering algorithm to group major sulcal patterns in 150 adult brains. However, the discrimination of 3D moment invariants was limited in distinguishing different patterns. Hence, a more representative descriptor was proposed in [4], where the distance between any two sulcal folds in 62 adult brains was computed after they were aligned, resulting in more meaningful results. Meanwhile, sulcal pits, the locally deepest points in cortical sulci, were proposed for studying the inter-individual variability of cortical folding [5]. This is because sulcal pits have been suggested to be genetically affected and closely related to functional areas [6]. It has been found that the spatial distribution of sulcal pits is relatively consistent across individuals, compared to the shallow folding regions, in both adults (148 subjects) and infants (73 subjects) [7, 8].

In this paper, we propose a novel method for discovering major representative patterns of cortical folding on a large-scale neonatal dataset ($N = 677$). The motivation of using a neonatal dataset is that all primary cortical folding is largely genetically determined and has been established at term birth [9]; hence, neonates with the minimal exposure to the complicated postnatal environmental influence are the ideal candidates

for discovering the major cortical patterns. This is very important for understanding the biological relationships between cortical folding and brain functional development or neurodevelopmental disorders rooted during infancy. The motivation of using a large-scale dataset is that small datasets may not sufficiently cover all kinds of major cortical patterns and thus would likely lead to biased results.

In our method, we leveraged the reliable deep sulcal pits to characterize the cortical folding, and thus eliminating the effects of noisy shallow folding regions that are extremely heterogeneous and variable. Specifically, first, sulcal pits were extracted using a watershed algorithm [8] and represented using a sulcal graph. Then, the difference between sulcal pit distributions of any two cortices was computed based on six complementary measurements, i.e., sulcal pit position, sulcal pit depth, ridge point depth, sulcal basin area, sulcal basin boundary, and sulcal pit local connection, thus resulting in six matrices. Next, these difference matrices were further converted to similarity matrices, and adaptively fused as one comprehensive similarity matrix using a similarity network fusion technique [10], to preserve their common information and also capture their complementary information. Finally, based on the fused similarity matrix, a hierarchical affinity propagation clustering algorithm was performed to group sulcal graphs into different clusters. The proposed method was applied to 677 neonatal brains (the largest neonatal dataset to our knowledge) in the central sulcus, superior temporal sulcus, and cingulate sulcus, and revealed multiple distinct and meaningful patterns of cortical folding in each region.

2 Methods

Subjects and Image Acquisition. MR images for $N = 677$ term-born neonates were acquired on a Siemens head-only 3T scanner with a circular polarized head coil. Before scanning, neonates were fed, swaddled, and fitted with ear protection. All neonates were unsexed during scanning. T1-weighted MR images with 160 axial slices were obtained using the parameters: TR = 1,820 ms, TE = 4.38 ms, and resolution = $1 \times 1 \times 1 \text{ mm}^3$. T2-weighted MR images with 70 axial slices were acquired with the parameters: TR = 7,380 ms, TE = 119 ms, and resolution = $1.25 \times 1.25 \times 1.95 \text{ mm}^3$.

Cortical Surface Mapping. All neonatal MRIs were processed using an infant-dedicated pipeline [2]. Specifically, it contained the steps of rigid alignment between T2 and T1 MR images, skull-stripping, intensity inhomogeneity correction, tissue segmentation, topology correction, cortical surface reconstruction, spherical mapping, spherical registration onto an infant surface atlas, and cortical surface resampling [2]. All results have been visually checked to ensure the quality.

Sulcal Pits Extraction and Sulcal Graph Construction. To characterize the sulcal folding patterns in each individual, sulcal pits, the locally deepest point of sulci, were extracted on each cortical surface (Fig. 1) using the method in [8]. The motivation is that deep sulcal pits were relatively consistent across individuals and stable during brain development as reported in [6], and thus were well suitable as reliable landmarks for characterizing sulcal folding. To exact sulcal pits, each cortical surface was

partitioned into small basins using a watershed method based on the sulcal depth map [11], and the deepest point of each basin was identified as a sulcal pit, after pruning noisy basins [8]. Then, a sulcal graph was constructed for each cortical surface as in [5]. Specifically, each sulcal pit was defined as a node, and two nodes were linked by an edge, if their corresponding basins were spatially connected.

Sulcal Graph Comparison. To compare two sulcal graphs, their similarities were measured using multiple metrics from spatial, geometrical, and topological points of view, to capture the multiple aspects of sulcal graphs. Specifically, we computed six distinct metrics, using sulcal pit position D , sulcal pit depth H , sulcal basin area S , sulcal basin boundary B , sulcal pit local connection C , and ridge point depth R . Given N sulcal graphs from N subjects, any two of them were compared using above six metrics, so a $N \times N$ matrix was constructed for each metric.

The difference between two sulcal graphs can be measured by comparing the attributes of the corresponding sulcal pits in the two graphs. In general, the difference between any sulcal-pit-wise attribute of sulcal graphs P and Q can be computed as

$$\text{Diff}(P, Q, \text{diff}_X) = \frac{1}{2} \left(\frac{1}{V_P} \sum_{i \in P} \text{diff}_X(i, Q) + \frac{1}{V_Q} \sum_{j \in Q} \text{diff}_X(j, P) \right) \quad (1)$$

where V_P and V_Q are respectively the numbers of sulcal pits in P and Q , and $\text{diff}_X(i, Q)$ is the difference of a specific attribute X between sulcal pit i and its corresponding sulcal pit in graph Q . Note that we treat the closest pit as the corresponding sulcal pit, as all cortical surfaces have been aligned to a spherical surface atlas.

(1) Sulcal Pit Position. Based on Eq. 1, the difference between P and Q in terms of sulcal pit positions is computed as $D(P, Q) = \text{Diff}(P, Q, \text{diff}_D)$, where $\text{diff}_D(i, Q)$ is the geodesic distance between sulcal pit i and its corresponding sulcal pit in Q on the spherical surface atlas.

(2) Sulcal Pit Depth. For each subject, the sulcal depth map is normalized by dividing by the maximum depth value, to reduce the effect of the brain size variation. The difference between P and Q in terms of sulcal pit depth is computed as $H(P, Q) = \text{Diff}(P, Q, \text{diff}_H)$, where $\text{diff}_H(i, Q)$ is the depth difference between sulcal pit i and its corresponding sulcal pit in Q .

(3) Sulcal Basin Area. To reduce the effect of surface area variation across subjects, the area of each basin is normalized by the area of the whole cortical surface. The difference between P and Q in terms of sulcal basin area of graphs P and Q is computed as $S(P, Q) = \text{Diff}(P, Q, \text{diff}_S)$, where $\text{diff}_S(i, Q)$ is the area difference between the basins of sulcal pit i and its corresponding sulcal pit in Q .

(4) Sulcal Basin Boundary. The difference between P and Q in terms of sulcal basin boundary is formulated as $B(P, Q) = \text{Diff}(P, Q, \text{diff}_B)$, where $\text{diff}_B(i, Q)$ is the difference between the sulcal basin boundaries of sulcal pit i and its corresponding sulcal pit in Q . Specifically, we define a vertex as a boundary vertex of a sulcal basin, if one of its neighboring vertices belongs to a different basin. Given two corresponding sulcal pits $i \in P$ and $i' \in Q$, their sulcal basin boundary vertices are respectively denoted as B_i and $B_{i'}$. For any boundary vertex $a \in B_i$, its closest vertex a' is found from $B_{i'}$; and similarly for any boundary vertex $b' \in B_{i'}$, its closest vertex b is found from B_i . Then,

the difference between the basin boundaries of sulcal pit i and its corresponding pit $i' \in Q$ is defined as:

$$\text{diff}_B(i, Q) = \frac{1}{2} \left(\frac{1}{N_{B_i}} \sum_{a \in B_i, a' \in B'_i} \text{dis}(a, a') + \frac{1}{N_{B'_i}} \sum_{b' \in B'_i, b \in B_i} \text{dis}(b', b) \right) \quad (2)$$

where N_{B_i} and $N_{B'_i}$ are respectively the numbers of vertices in B_i and B'_i , and $\text{dis}(\cdot)$ is the geodesic distance between two vertices on the spherical surface atlas.

(5) Sulcal Pit Local Connection. The difference between local connections of two graphs P and Q is computed as $C(P, Q) = \text{Diff}(P, Q, \text{diff}_C)$, where $\text{diff}_C(i, Q)$ is the difference of local connection after mapping sulcal pit i to graph Q . Specifically, for a sulcal pit i , assume k is one of its connected sulcal pits. Their corresponding sulcal pits in graph Q are respectively i' and k' . The change of local connection after mapping sulcal pit i to graph Q is measured by:

$$\text{diff}_C(i, Q) = \frac{1}{N_{G_i}} \sum_{k \in G_i} |\text{dis}(i, k) - \text{dis}(i', k')| \quad (3)$$

where G_i is the set of sulcal pits connecting to i , and N_{G_i} is the number of pits in G_i .

(6) Ridge Point Depth. Ridge points are the locations, where two sulcal basins meet. As suggested by [5], the depth of the ridge point is an important indicator for distinguishing sulcal patterns. Thus, we compute the difference between the average ridge point depth of sulcal graphs P and Q , as:

$$R(P, Q) = \left| \frac{1}{E_P} \sum_{e \in P} r_e - \frac{1}{E_Q} \sum_{e \in Q} r_e \right| \quad (4)$$

where E_P and E_Q are respectively the numbers of edges in P and Q ; e is the edge connecting two sulcal pits; and r_e is the normalized depth of ridge point in the edge e .

Sulcal Graph Similarity Fusion. The above six metrics measured the inter-individual differences of sulcal graphs from different points of view, and each provided complementary information to the others. To capture both the common information and the complementary information, we employed a similarity network fusion (SNF) method [10] to adaptively integrate all six metrics together. To do this, each difference matrix was normalized by its maximum elements, and then transformed into a similarity matrix as:

$$W_M(x, y) = \exp\left(-\frac{M^2(x, y)}{\mu \left(\frac{\Phi_x + \Phi_y + M(x, y)}{3}\right)}\right) \quad (5)$$

where μ was a scaling parameter; M could be any one of the above six matrices; Φ_x and Φ_y were respectively the average values of the smallest K elements in the x -th row and y -th row of M . Finally, six similarity matrices $W_D, W_H, W_R, W_S, W_B,$ and W_C were fused

together as a single similarity matrix W by using SNF with t iterations. The parameters were set as $\mu = 0.8$, $K = 30$, and $t = 20$ as suggested in [10].

Sulcal Pattern Clustering. To cluster sulcal graphs into different groups based on the fused similarity matrix W , we employed the Affinity Propagation Clustering (APC) algorithm [12], which could automatically determine the number of clusters based on the natural characteristics of data. However, since sulcal folding patterns were extremely variable across individuals, too many clusters were identified after performing APC, making it difficult to observe the most important major patterns. Therefore, we proposed a hierarchical APC framework to further group the clusters. Specifically, after running APC, (1) the exemplars of all clusters were used to perform a new-level APC, so less clusters were generated. Since the old clusters were merged, the old exemplars may be no longer representative for the new clusters. Thus, (2) a new exemplar was selected for each cluster based on the maximal average similarity to all the other samples in the cluster. We repeated these steps, until the cluster number reduced to an expected level (<5).

3 Results

We extracted sulcal pits on cortical surfaces from 677 neonatal brains. To demonstrate the validity of our methods for discovering the cortical folding patterns, we employed three representative cortical regions, i.e., the central sulcus, superior temporal sulcus, and cingulate sulcus. For each cortical region, a 677×677 similarity matrix was computed using SNF and all subjects were then clustered into different groups by the hierarchical APC. To better explore the major folding patterns, an average cortical surface was constructed for each cluster, based on 20 representative cortical surfaces that are most similar to the exemplar in each cluster. All sulcal pits in each cluster were mapped onto the average surfaces.

For the **central sulcus**, three distinct folding patterns were identified, as shown in Fig. 2. In the pattern (a), two sulcal pits concentration areas can be observed, indicating two sulcal basins in the central sulcus. This pattern was further confirmed by six representative examples of individual subjects (second to seventh columns). In the pattern (b), three distinct sulcal pits concentration areas can be observed, with one extra area (basin 3) located in the most inferior portion of the central sulcus, compared to the pattern (a). In the pattern (c), three distinct sulcal pits concentration areas can be observed as in the pattern (b), but they are more concentrated. This is also confirmed by six representative examples of (c). Moreover, compared to the pattern (b), the sulcal basin 2 is very short, while the sulcal basin 3 is very long in the pattern (c). Such phenomenon is likely related to “hand knob shift” in a study of the shape of the central sulcus in adults [13]. Previously, different studies reported either two [8] or three [7] sulcal basins in the central sulcus. Herein, we can see that both two-basin and three-basin patterns are major patterns of sulcal folding.

For the **superior temporal sulcus (STS)**, three distinct folding patterns were identified, as shown in Fig. 3. In the pattern (a), the distribution of sulcal pits in the posterior portion of STS is more diffused and bended, compared to the patterns (b) and

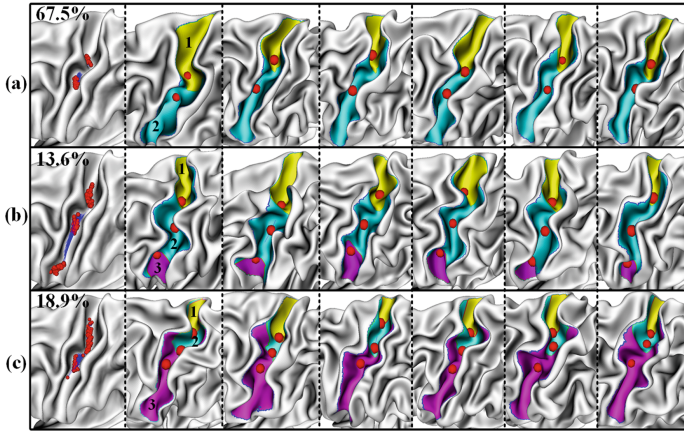


Fig. 2. Sulcal folding patterns in the central sulcus. The first column shows three discovered sulcal folding patterns, with all sulcal pits (red spheres) mapped onto the average surface of each cluster. For each pattern, the second to seventh columns show six representative examples of individual subjects. Different sulcal basins are marked with different colors. The percentage of each pattern is shown at the top-left corner.

(c), indicating the differences in the folding shape of STS. This is supported by a previous cortical folding study in adults, which reported that for some brains there was a Y-shaped STS but for some brains there was a single long STS [4]. In the pattern (b), compared to (a) and (c), an extra concentration region of sulcal pits is exhibited near the temporal pole, which is also confirmed by six representative examples in individual subjects, showing small sulcal basins near the temporal pole. In the pattern (c), the sulcal basin in the anterior portion of STS is very long and straight, extending to the temporal pole.

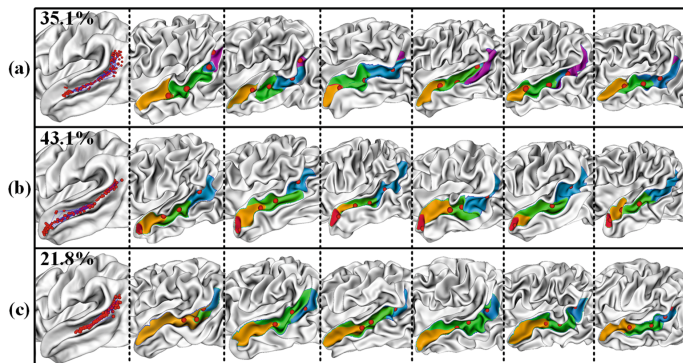


Fig. 3. Sulcal folding patterns in the superior temporal sulcus. The first column shows three discovered sulcal folding patterns, with all sulcal pits (red spheres) mapped onto the average surface of each cluster. For each pattern, the second to seventh columns show six representative examples of individual subjects. Different sulcal basins are marked with different colors.

For the **cingulate sulcus**, four distinct major folding patterns were identified, as shown in Fig. 4. In the pattern (a), a single long cingulate sulcus is clearly shown, while in the pattern (b), two long parallel sulci are observed. This is consistent with the previous cortical folding pattern study in adults [4], which reported that two cingulate sulci were observed in some brains. A study of autopsy specimen brains also reported that 24 % left hemispheres had double parallel cingulate sulcus [1]. In the pattern (c), the cingulate sulcus is interrupted in the anterior region; in contrast, in the pattern (d), the cingulate sulcus is interrupted in the posterior region. This two types of interruption were also reported in [1]. In pattern (c) and pattern (d), some parallel sulci can be observed, but they are much shorter than that in pattern (b).

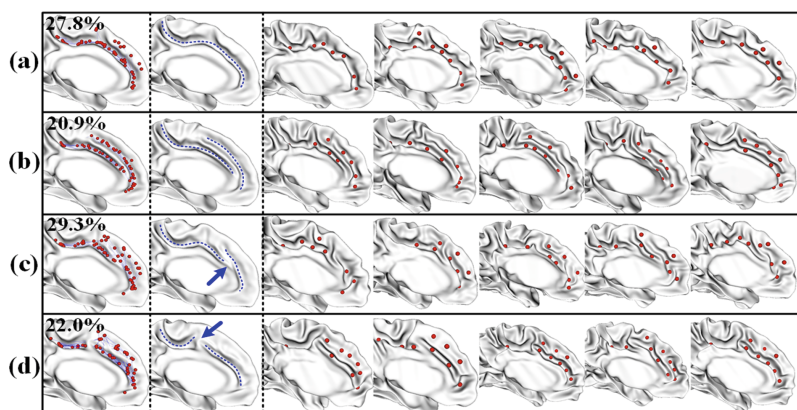


Fig. 4. Sulcal folding patterns in the cingulate sulcus. The first column shows four discovered folding patterns, with all sulcal pits (red spheres) mapped onto the average surface of each cluster. The second column shows the schematic drawing of the sulcal curves (blue dashes) on the average surface of each cluster. For each pattern, the third to seventh columns show five representative examples of individual subjects. The percentage of each pattern is shown at the top-left corner.

4 Conclusion

The main contribution of this paper is twofold. *First*, a novel generic method for discovering the cortical folding patterns was proposed, by leveraging the reliable sulcal pits. Specifically, multiple complementary similarity measures of sulcal pits graph were first computed and adaptively fused to comprehensively capture the individual similarity. Then, based on the fused similarity, sulcal pits graphs were clustered using a hierarchical affinity propagation algorithm. *Second*, for the first time, we applied the proposed method to discover the cortical folding patterns in a large-scale neonatal dataset with **677** subjects, and revealed multiple distinct and representative patterns. These results suggested that it is needed to construct multiple representative cortical folding atlases for each region for better spatial normalization of individuals in

group-level studies. Our future work includes discovering patterns in other cortical regions, and exploring their relationships with structural connectivity and cognitive functions.

Acknowledgements. This work was supported in part by UNC BRIC-Radiology start-up fund and NIH grants (MH107815, MH108914, MH100217, HD053000, and MH070890).

References

1. Ono, M., Kubik, S., Abernathy, C.D.: Atlas of the Cerebral Sulci. Thieme, New York (1990)
2. Li, G., Wang, L., Shi, F., et al.: Construction of 4D high-definition cortical surface atlases of infants: methods and applications. *Med. Image Anal.* **25**, 22–36 (2015)
3. Sun, Z.Y., Rivière, D., Poupon, F., Régis, J., Mangin, J.-F.: Automatic inference of sulcus patterns using 3D moment invariants. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part I. LNCS, vol. 4791, pp. 515–522. Springer, Heidelberg (2007)
4. Sun, Z.Y., Perrot, M., Tucholka, A., Rivière, D., Mangin, J.-F.: Constructing a dictionary of human brain folding patterns. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009, Part II. LNCS, vol. 5762, pp. 117–124. Springer, Heidelberg (2009)
5. Im, K., Raschle, N.M., Smith, S.A., et al.: Atypical sulcal pattern in children with developmental dyslexia and at-risk kindergarteners. *Cereb. Cortex* **26**, 1138–1148 (2016)
6. Lohmann, G., von Cramon, D.Y., Colchester, A.C.: Deep sulcal landmarks provide an organizing framework for human cortical folding. *Cereb. Cortex* **18**, 1415–1420 (2008)
7. Im, K., Jo, H.J., Mangin, J.F., et al.: Spatial distribution of deep sulcal landmarks and hemispherical asymmetry on the cortical surface. *Cereb. Cortex* **20**, 602–611 (2010)
8. Meng, Y., Li, G., Lin, W., et al.: Spatial distribution and longitudinal development of deep cortical sulcal landmarks in infants. *NeuroImage* **100**, 206–218 (2014)
9. Li, G., Nie, J., Wang, L., et al.: Mapping region-specific longitudinal cortical surface expansion from birth to 2 years of age. *Cereb. Cortex* **23**, 2724–2733 (2013)
10. Wang, B., Mezlini, A.M., Demir, F., et al.: Similarity network fusion for aggregating data types on a genomic scale. *Nat. Methods* **11**, 333–337 (2014)
11. Li, G., Nie, J., Wang, L., et al.: Mapping longitudinal hemispheric structural asymmetries of the human cerebral cortex from birth to 2 years of age. *Cereb. Cortex* **24**, 1289–1300 (2014)
12. Frey, B.J., Dueck, D.: Clustering by passing messages between data points. *Science* **315**, 972–976 (2007)
13. Sun, Z.Y., Kloppel, S., Riviere, D., et al.: The effect of handedness on the shape of the central sulcus. *NeuroImage* **60**, 332–339 (2012)