Genetic Effects on Fine-Grained Human Cortical Regionalization

Yue Cui1,2†, Bing Liu1,2†, Yuan Zhou4,5†, Lingzhong Fan1,2, Jin Li1,2, Yun Zhang6, Huawang Wu6,7, Bing Hou1,2, Chao Wang6, Fanfan Zheng1,2, Chengxiang Qiu1,2, Li-Lin Rao4,5, Yuping Ning7, Shu Li4,5 and Tianzi Jiang1,2,3,6,8

1Brainnetome Center, 2National Laboratory of Pattern Recognition, 3CAS Center for Excellence in Brain Science, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China, 4Key Laboratory of Behavioral Science, 5Magnetic Resonance Imaging Research Center, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China, 6Key Laboratory for NeuroInformation of Ministry of Education, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 610054, China, 7Guangzhou Brain Hospital, The Affiliated Brain Hospital of Guangzhou Medical University, Guangzhou 510370, China and 8Queensland Brain Institute, University of Queensland, Brisbane QLD 4072, Australia

Address correspondence to Tianzi Jiang, Brainnetome Center, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China. Email: jiangtz@nlpr.ia.ac.cn; Shu Li, Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China. Email: lishu@psych.ac.cn

†Y.C., B.L., and Y.Z. contributed equally to this work.

Abstract

Various brain structural and functional features such as cytoarchitecture, topographic mapping, gyral/sulcal anatomy, and anatomical and functional connectivity have been used in human brain parcellation. However, the fine-grained intrinsic genetic architecture of the cortex remains unknown. In the present study, we parcellated specific regions of the cortex into subregions based on genetic correlations (i.e., shared genetic influences) between the surface area of each pair of cortical locations within the seed region. The genetic correlations were estimated by comparing the correlations of the surface area between monozygotic and dizygotic twins using bivariate twin models. Our genetic subdivisions of diverse brain regions were reproducible across 2 independent datasets and corresponded closely to fine-grained functional specializations. Furthermore, subregional genetic correlation profiles were generally consistent with functional connectivity patterns. Our findings indicate that the magnitude of the genetic covariance in brain anatomy could be used to delineate the boundaries of functional subregions of the brain and may be of value in the next generation human brain atlas.

Key words: cortical regionalization, genetic correlation, genetics, surface area, twins

Introduction

Mapping fine-grained, anatomically distinct, and functionally specialized cortical subregions is fundamental for understanding brain function. Anatomical microstructure is currently primarily used to define cortical boundaries, and cyto-, myelo-, and receptor-architectonic maps have become the "gold standard" for cortical parcellation (Brodmann 1909; Amunts et al. 2013). In addition, many other techniques for parcellating the human brain, such as topographic mapping (Wandell and Winawer 2011), gyral/sulcal anatomy (Van Essen et al. 2012), and anatomical (Behrens et al. 2003) and functional (Kim et al. 2010) connectivity with in vivo magnetic resonance imaging (MRI), have been explored. These approaches used various brain imaging measures...
(i.e., phenotypes) that are unique to specific subregions. Evidence has suggested that these phenotypes are under genetic control (Karlsodt et al. 2010; Blokland et al. 2011; Chiang et al. 2011; Blokland et al. 2012). However, whether genetic information can feasibly be used to identify fine-grained cortical subregions and reveal the genetic basis of cortical regionalization is unknown.

The formation of brain phenotypes that show subregional specialization has generally been believed to be strictly controlled by molecular cascades. For example, the precise development of the cortical circuitry of fiber tracts involves families of molecules that direct the growth of axons (Dye et al. 2011) as well as axonal path finding (Molnar and Blakemore 1995; Chedotal and Richards 2010) and target selection (Ma et al. 2002). Prior to axonal guidance, however, differential gene expression plays an important role in the process of cortical regionalization in the early brain (Bishop et al. 2000; Sur and Rubenstein 2005; O’Leary et al. 2007). This occurs across the whole brain and in subregions such as the frontal cortex (Cholfin and Rubenstein 2007, 2008). In addition, the results of parcellations based on various structural and functional features are consistent for some regions (Van Essen and Glasser 2014), strongly implying that the distinguishable phenotypes may be influenced by an underlying genetic mechanism. Taken together, accumulating evidence suggests that the influence of genes may be vitally important for the high level of specialization in cortical regions.

Although each phenotype is related to a number of genes and their complex interactions, twin studies permit the investigation of the overall contributions of genetic factors to the variance of these phenotypes by comparing correlations between monozygotic (MZ) twins, who share 100% of their genes, and dizygotic (DZ) twins, who share ~50% of their genes. This method can also be used to examine genetic correlations, which represent the shared genetic influences of different points of surface area on cortical surface. A correlation coefficient of 1 indicates a perfect positive relationship, suggesting that the 2 points of surface area are influenced by common genes (i.e., share a common genetic basis). A correlation coefficient of 0 shows that there is no relationship present, suggesting that the 2 points of surface area are influenced by totally different genes. Given that genetic factors play an important role in the process of brain development and cortical patterning, we hypothesized that the genetic mechanisms that underlie cortical segregation may be reflected by genetic correlations. These could account for the magnitude of the genetic covariation in brain anatomy at various cortical locations and be used to delineate functional boundaries in the cortex. To test our hypothesis, we used classical twin analysis (Neale and Cardon 1992) to detect genetic correlations between various locations of the cortical surface area in a noninvasive manner. In the present study, we collected 2 independent brain imaging datasets on healthy young twins from different parts of China. Maps of the genetic correlations were created between each vertex and all other locations within the seed regions of interest for the purpose of dividing them into subregions. We also explored the differences between the genetic correlation profiles of the various subregions by using genetic correlations between the locations in the seed region and all locations across the entire brain hemisphere.

Materials and Methods

Participants

Discovery Dataset

The discovery dataset included a total of 222 healthy young Chinese same-sex twins from the Beijing Twin Study (BeTwiSt) of the Institute of Psychology, Chinese Academy of Sciences (Chen et al. 2013). The exclusion of an individual with incomplete scanning, an individual with excessive head motion, and their co-twins resulted in 218 participants (109 pairs) comprising 124 MZ and 94 DZ individuals (mean age, 19.0 ± 1.5 years; 62 male MZ, 62 female MZ, 48 male DZ, and 46 female DZ; all twins were complete twins). None of the participants had a history of psychiatric diagnoses or neurological or metabolic illnesses. They all gave full written informed consent to participate in the study. This study was approved by the Institutional Review Board of the Institute of Psychology of the Chinese Academy of Sciences and the Institutional Review Board of the Beijing MRI Centre for Brain Research. The zygosity of the 109 pairs of same-sex twins was determined by DNA analyses and by a questionnaire method (Chen et al. 2010). Of the 109 pairs of same-sex twins, 104 pairs were determined by DNA analyses. The classification accuracy of the DNA analysis approximated 100%. The remaining 5 pairs, whose saliva samples yielded insufficient DNA, were determined by a combination of parent reports and children’s self-reports about co-twin physical similarities and the frequency of confusion. This questionnaire method held a predictive accuracy of 91% (Chen et al. 2010).

Replication Dataset

To validate the reliability of the genetically based parcellation, we also conducted our parcellation scheme on an independent data set. The replication dataset was obtained from another cohort of young Chinese twins. At the time of this analysis, a total of 191 healthy young twins who, together with their first-, second-, and third-degree relatives, had no history of psychotic disorder had been scanned. We randomly excluded a triplet from a set of same-sex trzygotic triplets to form a twin pair. The resulting sample included 190 participants comprising 108 MZ and 82 DZ (mean age, 17.9 ± 2.7 years; 44 male MZ, 64 female MZ, 10 male DZ, 44 female DZ, and 28 opposite-sex pairs included in DZ twins; all twins were complete twins). All subjects gave full written informed consent to participate in the study. This study was approved by the Institutional Review Board of the Guangzhou Brain Hospital. The zygosity of each of the same-sex twin pairs was determined using 16 multiplex STR markers (PowerPlex 16 system; Promega) at the Forensic Medicine Department of Sun Yat-Sen University. Opposite-sex twin pairs were deemed to be dizygotic without genotyping.

Image Acquisition and Processing

Images from the discovery dataset were acquired with a 3.0 T Siemens TrioTim scanner. A three-dimensional T1-weighted volumetric sequence was performed using a protocol with repetition time (TR) = 2530 ms, echo time (TE) = 3.37 ms, flip angle (FA) = 7°, field of view (FOV) = 256 × 256 mm², acquisition matrix = 256 × 192, slice thickness = 1.33 mm without gap and slice number = 144. Resting-state fMRI data were obtained with an echo-planar imaging sequence with the following parameters: TE = 30 ms, TR = 2 s, FA = 90°, acquisition matrix = 64 × 64, FOV = 220 × 220 mm², slice thickness = 4 mm without gap, 32 transversal slices, and 180 volumes. The replication dataset was acquired on a different MRI machine with different scanning parameters. Specifically, the images were acquired with a 3.0 T Philips Achieva scanner. A three-dimensional T1-weighted volumetric sequence was performed using a protocol with TR = 8.25 ms, TE = 3.79 ms, FA = 7°, FOV = 256 × 256 mm², acquisition matrix = 256 × 256, slice thickness = 1 mm without gap, and slice number = 188.
The cortical surface reconstruction used the publicly available FreeSurfer software package, version 5.3.0 (http://surfer.nmr.mgh.harvard.edu/). The details of the processing techniques have been described elsewhere (Dale et al. 1999; Fischl et al. 1999; Fischl and Dale 2000). Briefly, the following stages were included: Talairach transformation, intensity inhomogeneity correction, removal of nonbrain tissues, intensity normalization, tissue segmentation, automated correction of topology defects, and surface deformation to form gray matter/white matter (white) and gray matter/cerebrospinal fluid (pial) boundary surface triangulations (Dale et al. 1999; Fischl and Dale 2000). The generated cortical surfaces were then carefully reviewed and manually edited for technical accuracy. Vertex-wise estimates of the surface area were calculated by assigning one-third of the area of each triangle to each of its vertices. We used 2819-iteration nearest-neighbor averaging to smooth the vertex-wise maps to ensure an accurate estimation of genetic correlations, as previously investigated by Chen and colleagues (Chen et al. 2012).

**Definition of Seed Regions**

Our genetically based parcellation scheme (Fig. 1) was investigated using representative cortical regions with evolutionary and functional diversity: the superior medial frontal cortex (SMFC), frontal pole (FP), inferior frontal gyrus (IFG), and primary motor cortex (M1). The seed SMFC was extracted from the AAL template and then projected to a cortical surface model. We used the same seed FP as our previous parcellation that used a connectivity-based technique; see Liu et al. (2013) for the details. In brief, the FP was manually edited on the basis of the FP extracted from the Harvard–Oxford cortical structure atlas. The volumetric FP was projected to a cortical surface to obtain a surface-based region. The IFG in the present study represented traditionally language-related regions, comprising the lateral triangular and opercular areas (pars opercularis and pars triangularis; approximately corresponding to Brodmann’s areas (BA) 44 and 45 [Nishitani et al. 2005]). This seed region was extracted from the FreeSurfer Desikan surface-based atlas (Desikan et al. 2006). The M1 seed region was obtained from the PALS Brodmann area atlas (Van Essen 2005), a cortical surface parcellation atlas included in the FreeSurfer software package.

**Twin Analysis**

In twin analyses, the variance of a phenotype can be accounted for by additive genetic influences (A), common environmental influences (C), or unique environmental influences, including measurement errors (E). A univariate model is able to estimate...
A, C, and E influences on the variance of the cortical surface area. However, based on previous findings, the cortical surface area shows little evidence of influences from C (Chen et al. 2011; Eyler et al. 2011). Therefore, we used a twin model that only contained A and E. The univariate AE model can be extended to a bivariate model, which can explain the sources of genetic and environmental covariance. Specifically, in addition to examining the genetic and environmental influences on the surface area of each vertex in the seed region, the bivariate related-factors model allows for estimates of the genetic ($r_G$) and environmental ($r_E$) correlations between the surface area of every pair of vertices (Neale and Cardon 1992). Here, $r_G$ represents the extent of overlap of the genetic factors influencing the surface area, and $r_E$ indicates the extent of the overlap of the environmental factors as well as measurement errors. The analyses were performed using the OpenMx package (version 1.3), a free structural equation modeling software integrated into the R environment (Boker et al. 2011). Before the model fitting, the vertex-wise surface area data were adjusted for age, sex, and global effects (i.e., the vertex-wise surface area at each vertex was divided by the total surface area).

**Genetic Correlation-Based Parcellation**

A genetic correlation map was generated by pairwise correlations between the vertices within the seed regions. These were first down-sampled from the original 163,842 to 10,242 vertices per hemisphere in order to reduce the computation time for genetic estimation. Given the level of smoothing of the cortical surface area, this down-sampling does not involve any significant loss of information. After obtaining a genetic correlation map, which consisted of the $r_{GH}$ between the surface area measurements for each pair of vertices, a spectral clustering algorithm was used for automatic clustering (Ng et al. 2001) of the left and right hemispheres, separately. Spectral clustering is an unsupervised machine learning algorithm that groups vertices that share similar genetic correlations. The optimal number of clusters was evaluated by the Calinski–Harabasz (CH) index (Calinski and Harabasz 1974), which calculates the ratio of the between-cluster ($A$) and the within-cluster variance ($B$): $CH = (A/B) \times ([N - k]/[k - 1])$, where $N$ is the number of observations and $k$ is the number of clusters. Generally, the optimal cluster number was chosen as the one that maximizes the CH index, that is, when $A$ is large and $B$ is small. Here, we synthesized the CH indices of the discovery and replication datasets and previous phenotypically based parcellation numbers to determine the optimal genetically based number of clusters for each seed region.

**Comparisons of Whole Hemisphere Genetic Correlations of Subregions**

To explore the specific genetic correlation profiles of the various subregions, comparisons of the genetic correlations between the subregions and the whole hemisphere were performed. Specifically, for each vertex in a subregion, we calculated its genetic correlations to all the vertices in the entire hemisphere, yielding a genetic correlation matrix $M_{v \times h}$, where $v$ is the number of vertices in a given subregion and $h$ is number of vertices in the corresponding hemisphere. We averaged each column of the matrix and obtained the subregion’s mean genetic correlation to every vertex in the hemisphere. Two-sample t-tests between every 2 subregions’ genetic correlations to all the vertices in the hemisphere were then performed to obtain the contrast maps. Random field theory was then used for a multiple comparisons correction (Hayasaka et al. 2004). The statistical analysis was performed using the SurfStat package (http://www.math.mcgill.ca/keith/surfstat/).

**Relationships between the Similarities in Genetic Correlation Profiles and Brain Connectivity Patterns**

We measured the consistency between 2 kinds of pairwise similarities among the vertices within the SMFC. The 2 kinds of similarities reflected genetic correlation profiles and functional connectivity patterns, both of which were calculated based on the correlations between the SMFC and the entire hemisphere. Diffusion tensor imaging was not carried out in the discovery dataset, which limits our measure of anatomical connectivity. However, functional connectivity with resting-state fMRI is also an important aspect of brain connectivity, which reflects brain regions sharing functional profiles. The processing pipeline is shown in Fig. 7A. For the genetic correlation analysis, the number of surface area measurements was down-sampled to 2562 vertices per hemisphere (as targets), of which 70 were in the left SMFC (as seeds). A genetic correlation matrix was generated by calculating the correlations between each seed and target ($O_{70 \times 2562}$). A genetic similarity matrix (dimensions: $70 \times 70$) was then obtained by calculating the similarity between each pair of seeds, that is, $W_G = G \times G^T$. In order to maintain the same numbers of seeds and targets for both the genetic correlation and the brain connectivity calculation, seed and target masks were generated from the projection of the surface-based vertices to the volume-based voxels using FreeSurfer. The preprocessing of the resting-state fMRI data was carried out using the DPARSF toolbox (http://www.restfmri.net/forum/DPARSF). The preprocessing steps included: (1) discarding the first 10 volumes of each functional time series to allow for magnetization equilibrium, (2) correcting the slice timing for the remaining 230 images and realigning them to the first volume for head motion correction (2 subjects were excluded either because the maximum displacement in the cardinal direction was $>3$ mm or because the maximum spin was $>3^\circ$), (3) spatial normalizing of all the volumes to the MNI EPI template, (4) spatial smoothing with a Gaussian kernel of 6 mm full-width at half maximum, (5) removing linear trends, (6) temporal band pass filtering (0.01–0.1 Hz), and (7) regressing out nuisance signals, such as those from the white matter and cerebrospinal fluid, as well as global signals and 6 motion parameters. Pearson’s correlations were performed between each seed and the target, generating a functional connectivity matrix ($F_{70 \times 2562}$). The similarity between each pair of seeds was then calculated as $W_F = F \times F^T$, resulting in a $70 \times 70$ functional cross-correlation matrix for each subject. We randomly selected 1 twin from each of the 101 pairs of twins with minimal head motion and averaged their functional similarity matrices (in total 101 matrices) to obtain an entry-wise mean functional similarity matrix. Finally, the Mantel test was used to quantify the relationship between the average functional similarity matrix and the genetic correlation similarity matrix. The Mantel test was conducted using the vegan package (Oksanen et al. 2015) in R.

**Results**

**Parcellation of the SMFC and Subregional Genetic Correlation Comparison**

As shown in Supplementary Figure 1, a cluster number of 2 was optimal for the SMFC. In agreement with existing structural- and functional-based parcellations, we found that the genetic
clustering of the SMFC showed anterior and posterior clusters that correspond to the presupplementary motor area (pre-SMA) and the SMA (Fig. 2A; for the re-arranged genetic correlation matrices of the seed regions, see Supplementary Fig. 2). Specifically, the anterior cluster was located just rostral to the vertical line extending from the anterior commissure (y ≈ 0), whereas the posterior cluster was located just caudal to the vertical line. The genetic architecture of the SMA and pre-SMA is in line with cytoarchitectonic (Zilles et al. 1996), anatomical (Johansen-Berg et al. 2004), and functional (Kim et al. 2010) connectivity-based parcellations. The parcellation topology was consistent between the replicate datasets (Supplementary Fig. 3).

We further investigated the genetic correlation profiles, which represent the genetic correlation distribution of the various subregions across the entire brain hemisphere (Fig. 2B). A two-sample independent t-test revealed significant differences between subregions (Fig. 2C). We found that the genetic correlations for the pre-SMA were significantly stronger with regions in the lateral prefrontal cortex and medial frontal cortex. In contrast, the significant correlations for the SMA were with regions in the motor and visual cortices. Fiber tractography and functional connectivity studies have shown that the pre-SMA has many projections to and from the prefrontal cortex and the anterior cingulate cortex (ACC), whereas the SMA is mainly connected to the nearby M1 (Johansen-Berg et al. 2004; Kim et al. 2010).

The subregional genetic correlation profiles were therefore largely consistent with their anatomical and functional connectivity patterns.

Parcellation of the IFG and Subregional Genetic Correlation Comparisons

In the case of the FP, the CH index indicated 3 or 4 solutions (Supplementary Fig. 1). The optimal number of clusters was previously investigated using a cross-validation technique in a connectivity-based parcellation scheme (Liu et al. 2013), and a cluster number of 3 was found to be the most appropriate value. This cluster number was therefore selected in the present study, which enabling a direct comparison of the genetic correlation- and anatomical connectivity-based divisions. We identified 3 separable subregions, FPo, FPm, and FPi, from the regional maps of the bilateral FP (Fig. 3A) using the genetic correlations within the FP. The left and right FP subregions presented similar patterns, which were consistent with the maximum probability maps provided by a connectivity-based parcellation with diffusion tensor imaging (Liu et al. 2013). The parcellation results were consistent between the replicate datasets (Supplementary Fig. 3).

Three subregions showed similar patterns with the highest correlation being with the frontal lobe, the lowest with the occipital lobe, and the median with the temporal and parietal lobes (Fig. 3B). Comparisons of the genetic correlation profiles between subregions (Fig. 3C) showed that the FPo had a stronger correlation with the temporal pole (TP) and the orbitofrontal cortex (OFC) compared with the FPi, as well as an increased trend with the supramarginal gyrus and the OFC compared with the FPm. The FPm showed a significantly increased correlation with the medial prefrontal cortex (MPFC) compared with the FPo. The FPi was found to have a stronger correlation with the dorsal prefrontal cortex (DPFC) compared with either the FPo or the FPm. A previous investigation based on white matter and functional connectivity found that the FP subregions are involved in distinct functional networks (Liu et al. 2013). In brief, the FPo anatomically connects with regions of the social emotion network including the OFC and TP, the FPm connects with areas of the default mode network including the ACC and MPFC, and the FPi connects with regions of the cognitive processing networks, including the DPFC. Therefore, the genetic correlation profiles of the various subregions were consistent with the connectivity patterns of the corresponding subregions.

Parcellation of the IFG and Subregional Genetic Correlation Comparison

As shown in Supplementary Figure 1, a cluster number of 5 reached a local peak for the IFG. We also used a parcellation number of 3 in order to test the resemblance to the cytoarchitectonic division. The IFG can be divided into BA 44 and 45 and the frontal operculum (FOP) in a three-cluster solution (Fig. 4A). Again, this finding was
largely in accordance with classical cytoarchitectonics, with a boundary that aligned with the diagonal sulcus (Nishitani et al. 2005). BA44 was subdivided into dorsal and ventral areas, 44d and 44v, and BA45 was subdivided into anterior and posterior areas, 45a and 45p, in the five-cluster solution (Fig. 4A), which corresponds to the subdivisions identified using transmitter receptor distribution data (Amunts et al. 2010).

BA44 and BA45 constitute the classical Broca’s area. Comparisons of the BA44 and BA45 subregions revealed that BA45 exhibited higher correlations with regions in the temporal lobe and prefrontal cortex compared with BA44. In contrast, for BA44, the genetic correlations were stronger with regions in the motor areas (Fig. 4C). The genetic correlation profiles of BA44 and 45 corresponded to the functional roles of language comprehension and speech production, respectively (Clos et al. 2013).

Parcellation of M1

For M1, a cluster number of 5 was found to be preferable because of the higher regional CH values in both datasets (Supplementary Fig. 1). M1 was able to be subdivided into 5 subregions, 5 of which corresponded to motor representations of body parts: the face, hand and arm, trunk, hip, and leg and foot (from ventrolateral to dorsomedial). The single remaining subregion in the anterior medial part of M1 was the SMA (Fig. 5). This is in line with widely recognized topographic organization. A connectivity-based parcellation of the lateral precentral gyrus resulted in 4 distinct
regions (Schubotz et al. 2010), in accordance with our parcellation of the lateral convexity of M1.

Ruling out Potential Underestimation from the Twin Models

To rule out the possibility that the power to detect genetic correlations might have been limited because of the fairly modest MRI twin samples, we combined the discovery and replication datasets to obtain a total of 204 pairs of healthy young twins. We used the same processing procedure as that used for the discovery dataset except that the site information was regressed out as another nuisance variable before the model fitting. The division of the seed regions displayed the same delineations as those obtained using the discovery and replication datasets separately (Fig. 6), demonstrating that our models are sufficient to detect genetic correlations from bivariate twin analyses.

Relationships Between the Similarities in Genetic Correlation Profiles and Brain Connectivity Patterns

To further characterize the quantitative relationships between the similarity in genetic correlation profile and the similarity in brain connectivity patterns, we measured resting state functional connectivity (using resting-state fMRI) in the discovery dataset. We found that the genetic correlation and functional connectivity similarity matrices were significantly correlated ($r = 0.344$, $P < 0.001$, Fig. 7B).

Discussion

To the best of our knowledge, this is the first study to parcellate fine-scale, functionally distinct subregions noninvasively based on intrinsic genetic information obtained by twin analysis. We found that the genetic clustering of the human SMFC, FP, IFG, and M1 was predominantly bilaterally symmetric and reproducible between 2 independent datasets. The resulting clusters closely resembled the subregions identified by other approaches using structural and functional features. Furthermore, we observed subregional genetic correlation profiles across each hemisphere which were generally consistent with previously identified structural and functional connectivity patterns. These results establish the presence of a close relationship between genetically driven and phenotypically driven cortical parcellations and further indicate that cortical functional segregation is primarily genetically determined.

Previous attempts at human brain parcellation have focused on the use of neuroimaging measures of brain structure and function. In the present study, we used a data-driven clustering approach.
algorithm that automatically grouped cortical locations that shared homogeneous genetic factors. Our identifications of the genetic borders were largely in agreement with previously defined structural and functional borders, indicating the validity of our parcellation scheme. Genetic factors participate in many aspects of cortical development, including neuron induction, polarization, migration, differentiation, and the intra- and inter-areal connections of nerve cells in the brain (Sanes et al. 2011), all of which are important for establishing and maintaining the regional identity of cells and tissues. For instance, reelin has been identified as a multifunctional protein that controls the positioning of neurons, together with their growth, maturation, and synaptic activity in the developing and adult brain (Lee et al. 2014). A number of transcription factors, such as Pax6 and Emx2, which are expressed in opposing gradients across the cortical surface, have also been shown to play an important role in the regional identities of cortical areas at the whole brain level (Bishop et al. 2000; O’Leary et al. 2007; Rakic et al. 2009). At the lobar level, evidence from animal studies using mutant mice supports the concept that genetic mechanisms govern the process of subdividing the frontal cortex (Cholfin and Rubenstein 2007, 2008). Specifically, frontal cortex subdivision patterning is regulated by Fgf8, Fgf17, and Emx2, which play distinct roles in molecular regionalization (Cholfin and Rubenstein 2008). Although many genes are likely to be involved in cortical regionalization, a particular set of genes may influence a specific
subregion, and different subregions may be controlled by distinct gene sets. High genetic correlations with respect to the surface area between any 2 surface locations indicate that they were controlled by a common gene set, and vertices with similar genetic correlations are considered to be located in the same subregion. This is the basic concept behind the idea that genetic correlations can provide the distinguishing features to parcellate subregions. Recently, Chen and colleagues investigated older human adult twins and demonstrated that genetic topography exhibited anterior–posterior and dorsal–ventral organizational gradients on a coarse level (Chen et al. 2011). This level of patterning by opposing gradients is supported by animal studies that used experimental inhibition or overexpression of specific genes (Bishop et al. 2000; O’Leary et al. 2007; Rakic et al. 2009). However, evidence of genetic influences on a finer scale has been sparse.

We observed a consistency between the similarities in genetic correlation profiles and in brain connectivity patterns. The evidence suggests that families of molecules regulate the development of the cortical circuitry, including the growth of axons (Dye et al. 2011), path finding (Molnar and Blakemore 1995), and target selection (Ma et al. 2002), which together impose cytoarchitectonic differentiation on the developing cortex. For instance, the growth of projections correlates with Lhx2, Id2, and Coup-Tf1 expression (Dye et al. 2013), and some guidance molecules have also been identified in the thalamocortical pathway (Ma et al. 2002). Brain function is constrained by fiber connectivity, which reflects the capacity for information flow within and between specific structures (Passingham et al. 2002). Because we found that the correlation between genetic correlation and brain connectivity similarity matrices is modest, it is likely that genetic correlations influence many other phenotypes in addition to brain connectivity.

The 4 regions examined in the present study are characterized by various phylogenetic and functional features. The SMFC is a phylogenetically older area, which is also part of the primate cerebral cortex. Its functional roles are clearly distinct, with the SMA phylogenetically older area, which is also part of the primate cerebral cortex. Its functional roles are clearly distinct, with the SMA being one of the phylogenetically youngest and most diverse regions in humans. Nevertheless, our results indicate that functional segregation can be detected using genetic information, suggesting a genetic origin for the functional subdivisions.

The IFG has long been associated with language functions, which are among the most distinctive attributes of humans. Our three-cluster solution for the genetically based parcellation of the IFG is congruent with classical cytoarchitectonically (Amunts et al. 1999) and connectivity-driven (Anwander et al. 2007) segregation schemes. BA 44 and 45, which have clear cytoarchitectonic differences, form the classical Broca’s area (Amunts et al. 1999). The segregation of the FOP and BA 44/45 agrees with previous findings that Broca’s area and the FOP are functionally and phylogenetically distinct (Friederici et al. 2006). Furthermore, our five-cluster solution suggested a further division within BA 44 and 45, which is consistent with more finely detailed multiple receptor mapping (Amunts et al. 2010), functional activation (Heim and Friederici 2003; Molnar-Szakacs et al. 2005), and connectivity-based studies (Neubert et al. 2014). In the present study, these 2 parcellation levels were independent of each other, providing powerful evidence of a hierarchical organization within BA 44 and 45. Indeed, genetically controlled hierarchical segregation has previously been found at the whole brain level (Chen et al. 2012). In addition, it has been suggested that the functional role of BA 44/45 is associated with complex and hierarchical structure processing (Friederici et al. 2006; Amunts et al. 2010). Notably, we found that BA 44 was more closely genetically correlated to the auditory cortex than was BA 44, a correlation that is supported by the known functional role of BA 44 in language comprehension (Clos et al. 2013). In contrast, we found that BA 44 was more closely genetically correlated to the motor area than was BA 44. Again, this relationship is supported by the known relationship of BA 44 with speech production (Clos et al. 2013). Although lateralization in structure and function has been found in the IFG, we identified similar genetically based parcellation topology in the left and right hemispheres. From a genetic point of view, symmetry between hemispheres is also a predominant feature on a whole brain level (Chen et al. 2012).

M1 is located at the posterior convolution of the frontal lobe. There is consensus that an orderly motor representation of the human body (also called motor somatotopy) exists in M1. We found that our genetic subdivisions corresponded to the gross organization of the motor homunculus. This is a typical region in which its functional coding is based on locations, but not on structural morphology, as a result of which the cytoarchitectonic division is not consistent with the topographic organization (Geyer et al. 1996). In general, connectivity-based approaches agree with functional topography. Nebel and colleagues have reported a functional connectivity analysis of M1, with a five-partition rough correspondence with the motor homunculus, albeit that the middle lateral section showed an anterior–posterior division (Nebel et al. 2014). Other tractography studies revealed a dorsal/ventral division (Tomassini et al. 2007) and detailed 4 subregions (Schubotz et al. 2010) of the lateral premotor cortex, approximately corresponding to the lateral portion of M1. Our parcellation of M1 indicated that the functional correspondence is primarily genetically influenced, suggesting a genetic basis for the functional subregions.

Taken together, the 4 regions examined in the present study provide a representative example of the use of the current phenotypically based parcellation approach incorporating evolutionary and functional variety. Our results indicate that genetic
correlation-based parcellation is reliable and robust and may be universal and generalizable to other regions of the human brain. Because MZ twins share all their genes whereas DZ twins share an average of 50% of theirs, bivariate twin models present a powerful strategy for estimating the degree of genetic overlap (genetic correlations) between phenotypes. In addition, the ability to assess genetic sources using MRI allows for an in vivo investigation of the basis of brain organization. We speculate that genetic correlations of cortical surface area of the human brain may reflect cortical regionalization that was established in early development, but further study is needed to verify this. Although our scanned twin sample was not extremely large, because our parcellation patterns were validated by an independent dataset and because the combination of the 2 datasets provided verification for 4 representative regions, we believe our findings to be credible and reliable. In addition, other seed regions and/or a more data-driven approach without predefined ROIs as well as the match between their genetic correlation profiles and functional connectivity patterns needs to be explored in future studies. It should also be noted that, although the spatial and temporal resolution of brain imaging in the present study is frequently used, it is not the highest available. In future studies, we will consider making use of the much higher quality publicly available data from the Human Connectome Project on this topic.

In conclusion, this study provides in vivo evidence that the genetic correlations of cortical surface area obtained by twin analysis can be used to identify fine-grained functional subregions. Our findings suggest that genetic correlations are generally interpretable by existing phenotypically based approaches, thereby having the potential to unravel population–based fundamental patterns of the cortex and of inter-regional connectivity. The present study is important for understanding the genetic basis of cortical regionalization and provides guidance and validation for the delineation of the next generation human brain atlas.

**Supplementary Material**

Supplementary Material can be found at http://www.cercor.oxfordjournals.org/ online.

**Funding**

This work was supported in part by the National Key Basic Research and Development Program (973) (Grant 2011CB707800), the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant XDB02030300), the Natural Science Foundation of China (Grants 91132301; 61305143; 31107976; 31300843), Youth Innovation Promotion Association, Chinese Academy of Science.

**Notes**

The authors thank Dr. Yu Luo for her technical assistance, Yun Wang, Dang Zheng, and Liqing Yang for their help in data collecting, Drs. Rowan Tweedale, Rhoda E., and Edmund F. Perozzi for English editing on the manuscript, and all the participating twins. The authors also thank the support from BeTwiSt of Institute of Psychology, Chinese Academy of Sciences. Conflict of Interest: None declared.

**References**


