Quantifying Differences and Similarities in Whole-Brain White Matter Architecture Using Local Connectome Fingerprints

Fang-Cheng Yeh[1*], Jean M. Vettel[2,3], Aarti Singh[4], Barnabas Poczos[4], Scott Grafton[3], Kirk I. Erickson[5], Wen-Yih I. Tseng[6], and Timothy D. Verstynen[1*]

[1]Department of Psychology and Center for the Neural Basis of Cognition, Carnegie Mellon University, Pennsylvania, USA.

[2]U.S. Army Research Laboratory, Aberdeen Proving Ground, Aberdeen, Maryland, USA.

[3]Department of Psychological and Brain Sciences, University of California, Santa Barbara, Santa Barbara, California, USA.

[4]Department of Machine Learning, Carnegie Mellon University, Pennsylvania, USA.

[5]Department of Psychology, University of Pittsburgh, Pittsburgh, Pennsylvania USA.

[6]Institute of Medical Device and Imaging and Molecular Imaging Center, National Taiwan University, Taipei, Taiwan.

*Correspondence to:

Timothy D. Verstynen Ph.D., Fang-Cheng Yeh Ph.D., M.D.

Department of Psychology and Center for the Neural Basis of Computation,

Carnegie Mellon University

Pittsburgh, PA, USA

Email: timothyv@andrew.cmu.edu, frankyeh@cmu.edu

*Keywords: Connectome, Connectomics, Connectome fingerprint, Diffusion MRI

Abstract

Quantifying differences or similarities in connectomes between individuals has been a challenge due to the immense complexity of global brain networks. Here we introduce a noninvasive method that uses diffusion MRI to characterize whole-brain white matter architecture as a single local connectome fingerprint that allows for a direct quantification of differences or similarities between two structural connectomes. In four independently acquired data sets with repeated scans (total N=213), we show that the local connectome fingerprint is highly specific to an individual, allowing for an accurate self-versus-others classification that achieves 100% accuracy in across 17,398 identification tests. The estimated classification error was approximately one thousand times smaller than fingerprints derived from fractional anisotropy or region-to-region connectivity patterns. We further illustrate that the local connectome fingerprint can be used as a phenotype, revealing 12.51% similarity between monozygotic twins, 5.14% between dizygotic twins, and 4.51% between none-twin siblings. Finally, we show that the local connectome fingerprint can quantify neuroplasticity over time as reflected by a decrease in self-similarity at an average rate of 7.26% per year. This novel quantification approach opens a new door for probing the influence of pathological, genetic, social, or environmental factors on the unique configuration of the human connectome.

*Keywords: Connectome, Connectomics, Connectome fingerprint, Diffusion MRI

Author Summary

We show that the local organization of white matter architecture is highly unique to individuals and reliable enough to the use of as a metric of brain difference. The difference in white matter architecture is found to be partially determined by genetic factors, but largely plastic across time. This approach opens a new door for probing the influence of pathological, genetic, social, or environmental factors on the unique configuration of the human connectome.

Introduction

The specific brain characteristics that define an individual are encoded by the unique pattern of connections between the billions of neurons in the brain [1]. This complex wiring system, termed the connectome [2, 3], reflects the architecture of the region-to-region connectivity [4] that may advance the understanding of what influences the white matter architecture. Yet to date, quantifying the difference between two brain structures is still a major challenge as it requires a reliable characterization of white matter architecture that reveals the microscopic features. To this end, studies have used diffusion MRI (dMRI) to measure the microscopic diffusion patterns of white matter structure and characterize its connectivity pattern [5, 6]. This also allowed for the mapping of the whiter matter trajectories in the human brain and defining the graph structure of region-to-region connectivity [7, 8]; however the efficiency and reliability of tractography approaches have recently come into question [9, 10]. Instead of mapping region-to-region connectivity, the concept of the *local* connectome has recently been proposed as an alternative measure of macroscopic white matter pathways that overcomes the limitations of diffusion MRI fiber tracking [9-11]. The local connectome is defined as the degree of connectivity between adjacent voxels within a white matter fascicle defined by the density of the diffusing water. A collection of these density measurements provides a high dimensional feature vector that can describe the unique configuration of the structural connectome within an individual, providing a novel approach for comparing differences and similarities between individuals as pairwise distances.

Here we use this local connectome feature vector as a fingerprint to quantify the similarity and difference between two white matter architectures. To evaluate the performance of our approach, we used four independently collected dMRI datasets (n=11, 25, 60, 118, see Methods) with repeat scans at different time intervals (ranging from the same day to a year) to examine whether local connectome fingerprint can reliably distinguish the difference between within subject and between subject scans. We examined whether this local connectome fingerprint is a unique identifier of an individual person by testing how reliable the fingerprint could determine whether two samples came from the same person or different individuals. This uniqueness was compared with fingerprints derived from fractional anisotropy (FA)[12] and conventional region-to-region connectivity methods. Follow-up analysis reveals on how local connectome fingerprints can

quantify the similarity between genetically-related individuals as well measure longitudinal changes in within an individual across time.

Results

Characterization of white matter architecture

We first illustrate the characterization of white matter architecture within an individual. Figure 1A shows the spin distribution functions (SDFs) [13] estimated from dMRI scans at the mid-sagittal section of the corpus callosum. SDF represents the density of water diffusing at any orientation, and its magnitudes at the axonal directions can be used as the density-based measurements to quantify the local connectome (see Methods). An example of the local connectome quantified at the corpus callosum is illustrated for three subjects in Fig. 1B. Here the anterior and posterior portion of corpus callosum exhibit substantial diversity between these three subjects. A repeat scan several months later reveals a qualitative within-subject consistency. This high individuality appears to be specific to diffusion density estimates, as conventional FA measures calculated from diffusivity do not yield this qualitative between-subject diversity (Fig. 1C). To sample the local density measurements across all major white matter pathways, dMRI data was reconstructed into a standard space, and the fiber directions of a common atlas were used to sample an SDF value for each fiber direction (see Methods and Fig. 2A). This approach yields, for each dMRI scan, a local connectome fingerprint consisting of a high-dimensional feature vector with a total of 513,316 density estimates (Fig. 2B). Fig. 2C shows the fingerprints of the same three subjects in Fig. 1B and the fingerprints from their repeat scans. Consistent with the qualitative measurements in Fig. 1B, each local connectome fingerprint in Fig. 2C shows, at a coarse level, a highly similar pattern for within-subject scans and also high variability across subjects, suggesting that the local connectome fingerprint may exhibit the unique features of the white matter architecture.

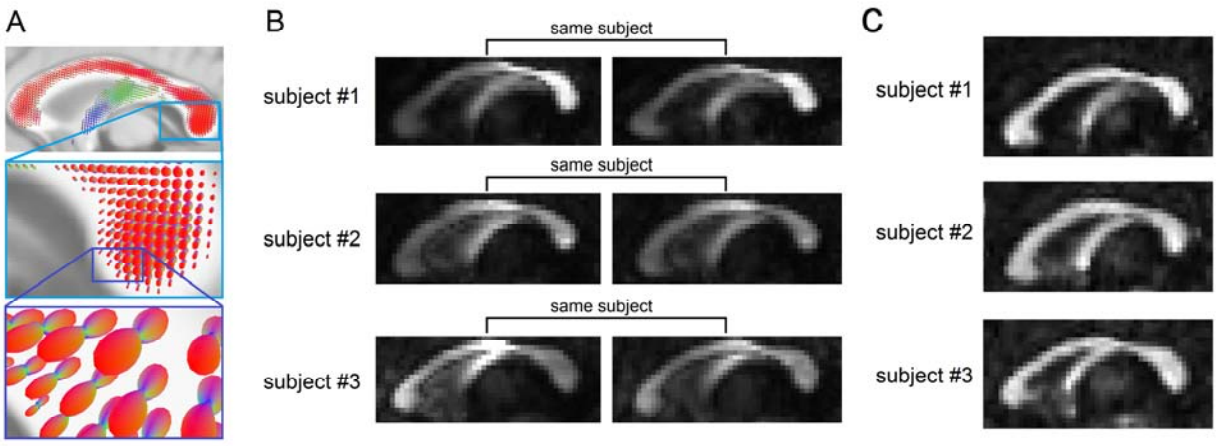

Fig 1. The uniqueness of local connectome structure revealed by the density of diffusing water.

(A) The spin distribution function (SDF) calculated from diffusion MRI quantifies the density of diffusing water along axonal fiber bundles. The magnitudes of SDF at axonal directions provide density-based measurements to characterize axonal fiber bundles. (B) The density measurements obtained from SDF show individuality between-subjects #1, #2, and #3 (intensity scaled between [0 0.8]). The density of diffusing water varies substantially across different portions of the corpus callosum. The repeat measurements after 238 (subject #1), 191 (subject #2), and 198 (subject #3) days still present a consistent pattern that captures individual variability. (C) In contrast to the SDF shown in (B), the fractional anisotropy derived from diffusivity shows no obvious individuality between the same subjects #1, #2, and #3 (intensity also scaled between [0 0.8]). This is due to the fact that diffusivity, which quantifies how *fast* water diffuses, does not vary a lot in normal axonal bundles.

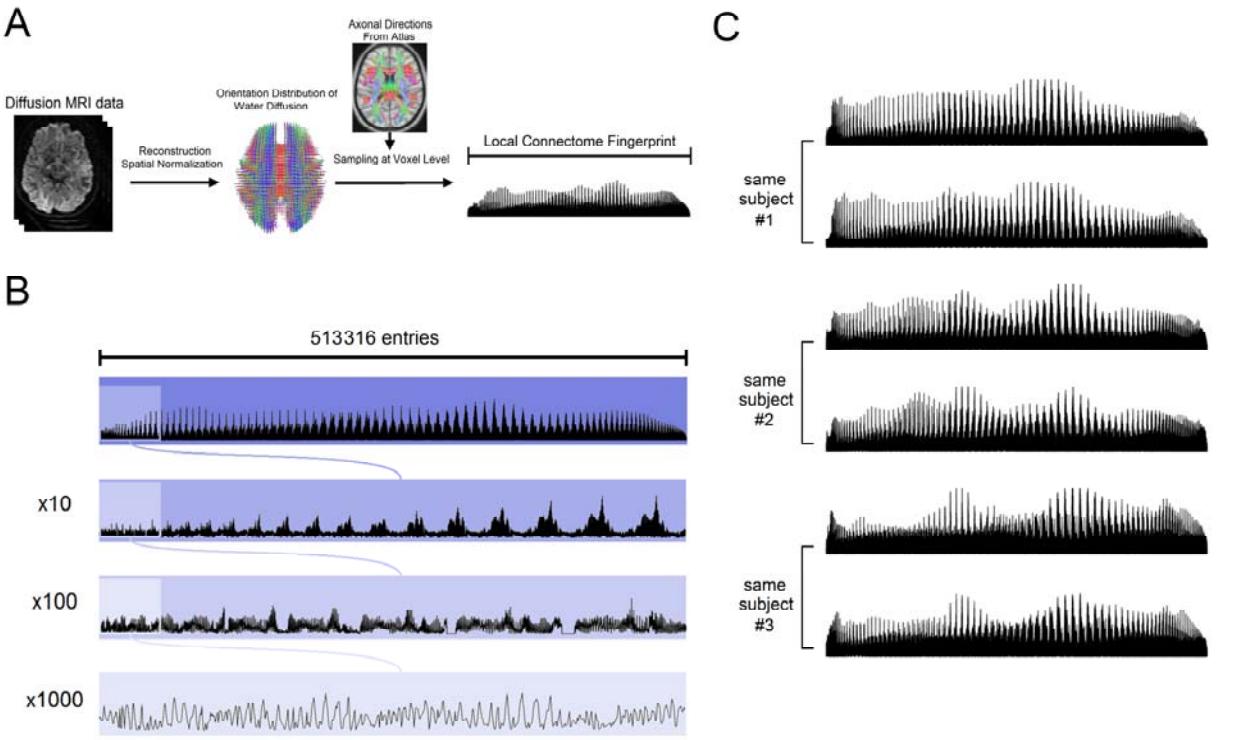

Fig 2. Local connectome fingerprinting.

(A) Local connectome fingerprinting is conducted by first reconstructing diffusion MRI data into a standard space to calculate the spin distribution functions (SDFs). A common fiber direction atlas is then used to sample the density of diffusing water along the principle directions in the cerebral white matter. The sampled measurements are compiled in a left-posterior-superior order to form a sequence of characteristic values as the local connectome fingerprint. (B) One local connectome fingerprint is shown in different zoom-in resolutions. A local connectome fingerprint has a total of 513,316 entries of scalar values presenting a unique pattern that can characterize the individuality of the human brain connections. (C) The local connectome fingerprint of subject #1, #2, and #3 and their repeat measurements (lower row) after 238, 191, and 198 days, respectively. At a coarse level, the local connectome fingerprint differs substantially between three subjects, whereas those from the repeat scans show a remarkably identical pattern, indicating the uniqueness and reproducibility of the local connectome fingerprint.

Between-subject versus within-subject difference

To quantify how well the local connectome fingerprint captures between-subject difference, we used four independently collected dMRI datasets (n=11, 24, 60, 118) with repeat scans for a subset of the subjects (n=11×3, 24×2, 14×2, 44×2, respectively). The Euclidian difference (i.e., root-mean-squared error) was used as a single difference estimate between any two fingerprints. For each dataset, we computed within-subject differences (n=33, 24, 14, 44, respectively) and between-subject differences (n=495, 1104, 2687, and 12,997, respectively). Figure 3 shows the within- and between-subject differences of the four datasets. All four datasets show a clear separation between the within-subject and between-subject difference distributions, with no single within-subject difference pairing being as large as any of the between-subject differences. The d-prime sensitivity index also shows high separability between the two difference distributions, with d-prime values of 14.84, 12.80, 7.21, and 8.12, for dataset I, II, III, and IV respectively.

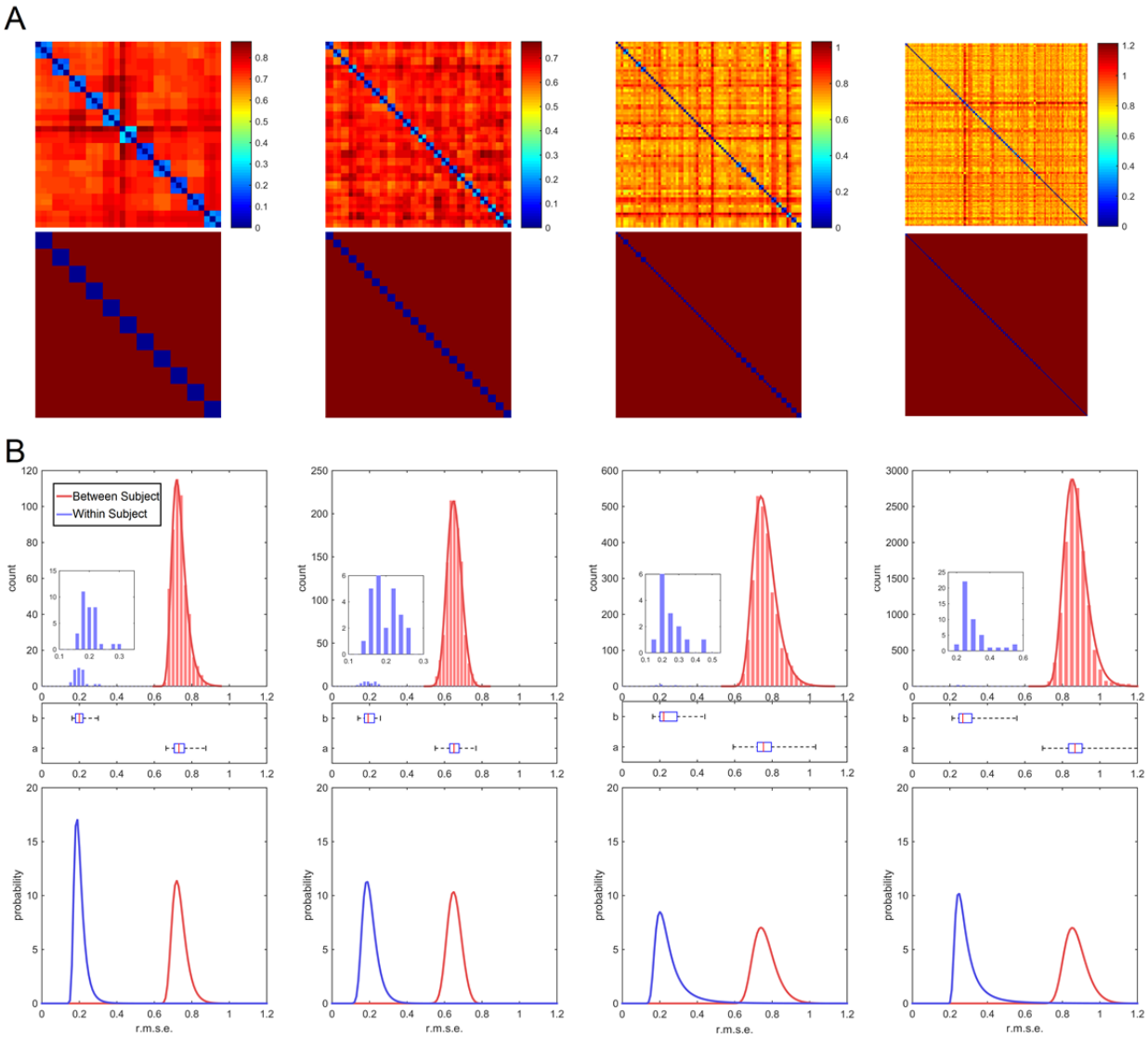

Fig 3. Within-subject versus between-subject difference in the local connectome fingerprints calculated from four datasets.

(A) The first row shows the matrix of pair-wise difference between any two local connectome fingerprints is calculated for dataset I, II, III, and IV (column 1, 2, 3, and 4). The second row shows the location of the within-subject (blue) and between-subject difference (red). The difference matrix shows substantially high between-subject difference and low within-subject difference. (B) The histograms of within-subject (blue) and between-subject (red) difference in the connectome fingerprints calculated from the dataset I, II, III, and IV (column 1, 2, 3, and 4). The first row shows the histograms, and the second row shows the box plot of their quartiles. In these four datasets, within-subject (blue) and between-subject (red) differences have perfect separation. In the last row, the histograms are fitted with generalized extreme value distribution

(also shown by solid curves in the second row) to estimate the classification error of the connectome fingerprint. The estimated classification error was 4.25×10^{-6}, 9.97×10^{-7}, 5.3×10^{-3}, and 5.5×10^{-3} for dataset I, II, III, and IV, respectively. The larger error in dataset III and IV could be due to their longer scanning interval (6 months and one year).

This consistency suggests that the local connectome fingerprint could be used as a unique characterization feature vector of an individual subject. To assess this, we used a linear discriminant analysis (LDA) classifier [14] to classify whether two fingerprints came from the same individual using only the difference between fingerprints as the classification feature. For each dataset, the classification error was estimated using leave-one-out cross-validation. Out of a total of 17,398 cross-validation rounds from four datasets (17,283 different-subject and 115 same-subject pairings), there was not a single misclassification. To estimate the true classification error, we modeled the distributions of within-subject and between-subject differences by the generalized extreme value distribution [15], a continuous probabilistic function often used to assess the probability of extreme values (smallest or largest) appearing in independent identically distributed random samples (last row of Fig. 3B). The classification error can be quantified by the probability of a within-subject difference greater than a between-subject difference. Our analysis shows that the classification error was 4.25×10^{-6} for dataset I, 9.97×10^{-7} for dataset II, 5.3×10^{-3} for dataset III, and 5.5×10^{-3} for dataset IV. The larger error in dataset III and IV could be due to their longer scan interval (6 months and one year). Thus, the probability of mistaking two samples of the same subject's local connectome fingerprint as coming from two different individuals is low enough so as to consider the local connectome fingerprint a highly reliable measure of individual subject uniqueness.

Corpus callosum fingerprint

Since the macroscopic patterns such as gyral and sulcal folding could also contribute to the characteristic features of the local connectome fingerprint, it is possible that the unique features we observed reflect only an artifact of the normalization process, not truly the microscopic characteristics of the white matter architecture. Therefore, we retested our measure within a restricted white matter mask that only covers the median sagittal sections of corpus callosum defined by the JHU white matter atlas. This "corpus callosum fingerprint" should be free from all

possible contribution of macroscopic pattern such as gyral and sulcal foldings. The same analysis procedures applied to the local connectome fingerprint were applied to the corpus callosum fingerprint to examine whether it reveal unique patterns specific to individuals. The result showed that the d-prime values were 5.97, 5.85, 3.78, and 4.08, for dataset I, II, III, and IV, respectively. The leave-one-out cross-validation analysis showed that classification error was 0%, 0.089%, 1.26%, and 0.63%, for dataset I, II, III, and IV, respectively. The classification error modeled by the generalized extreme value distribution was 9.13×10^{-4}, 5.6×10^{-3}, 6.9×10^{-3}, and 7.2×10^{-3}, for dataset I, II, III, and IV, respectively. The corpus callosum fingerprint itself already can achieve more than 99% accuracy in subject identification. This suggests that the high individuality of the local connectome fingerprint is not due macroscopic patterns such as gyral and sulcal folding but owing to the microscopic characteristics of the white matter architecture.

Comparison with FA-based fingerprints

FA has also been shown to reveal microscopic integrity of the axonal structure, and whether the density measure used in local connectome fingerprint can provide better characterization of white matter architecture requires a comparison. To this end, we used the identical analysis pipeline as was used for the local connectome fingerprint and only replaced the SDF-based measures with FA values of the corresponding voxels. Our analysis showed that the d-prime values of the FA-based fingerprint were 4.84, 4.70, 4.56, and 3.60, for dataset I, II, III, and IV, respectively. All values are substantially smaller than those of the local connectome fingerprint. The leave-one-out cross-validation analysis showed that classification error of the FA-based fingerprint was 0%, 0.18%, 0.22%, and 0.87%. While FA-based fingerprints also have high uniqueness with classification error less than 1%, it is clear from these results that the greatest reliability at characterizing connectomic uniqueness comes from the local connectome measures. This also indicates that SDF can provide better characterization of white matter architecture.

Comparison with global connectivity-based fingerprints

We further compared the local connectome fingerprint with region-to-region connectivity estimates from diffusion MRI fiber tracking. The same analysis pipeline as was used for the local connectome fingerprint was used to calculate leave-one-out cross-validation error for the traditional connectivity matrix. The d-prime values for the region-to-region connectivity matrices

in dataset I, II, III, and IV were at 3.44, 2.06, 2.41, and 2.25, respectively. The classification error for datasets I, II, III, and IV were 3.6%, 13.65%, 11.81%, and 9.48%, respectively (estimated by leave-one-out cross validation). While the classification accuracy for the traditional connectivity matrices is still quite high and similar to what has previously been observed in resting state functional connectivity estimates [16], it is clear from these results that the greatest reliability at characterizing connectomic uniqueness comes from local connectome measures.

Similarity among genetically-related individuals

The local connectome fingerprint opens the possibility for comparing not only differences but also the similarities between individuals, but it is crucial to examine whether it can reliably quantify white matter architecture as phenotypes. To this end, we used a publicly available dMRI dataset of 486 subjects from Human Connectome Project (2014, Q3 release), including 49 pairs of monozygotic (MZ) twins, 43 pairs of dizygotic twins (DZ) twins, and 96 pairs of non-twin siblings. While the local connectome fingerprints of MZ twins show generally similar patterns at the coarse level (Fig. 4), there are also substantial individual differences between the twins that can still be observed along the fingerprints. Consistent with these qualitative comparisons, we found that MZ twins have smaller differences between fingerprints, followed by DZ twins, siblings, and unrelated subjects (Fig. 5A). It is noteworthy that all difference distributions have large overlapping regions (Fig. 5B), indicating that the difference between twins or siblings may often fall within the distribution of differences from genetically-unrelated subjects. We further compared the similarity between twins and siblings. The similarity index of two local connectome fingerprints was calculated as a percentage of the mean difference between unrelated subjects (see Methods). On average, MZ twins have a similarity index of 12.51±1.09%, whereas similarity for DZ twins and siblings is 5.14±1.34% and 4.47±0.59%, respectively (Fig. 5C). The difference in similarity index was significant across MZ twins, DZ twins, non-twin siblings, and genetically-unrelated subjects ($F[3,51224] = 93.64$, $p < 0.001$). Post-hoc comparisons using Tukey-Kramer tests showed significant differences between any two groups (all $p < 0.001$), except between DZ twins and non-twin siblings ($p = 0.9348$). This result is consistent with our understanding that MZ twin shares higher genetic similarity, whereas DZ

twins share a similar genetic similarity as non-twin siblings. This suggests that local connectome fingerprint can be used as a reliable phenotype of white matter architecture.

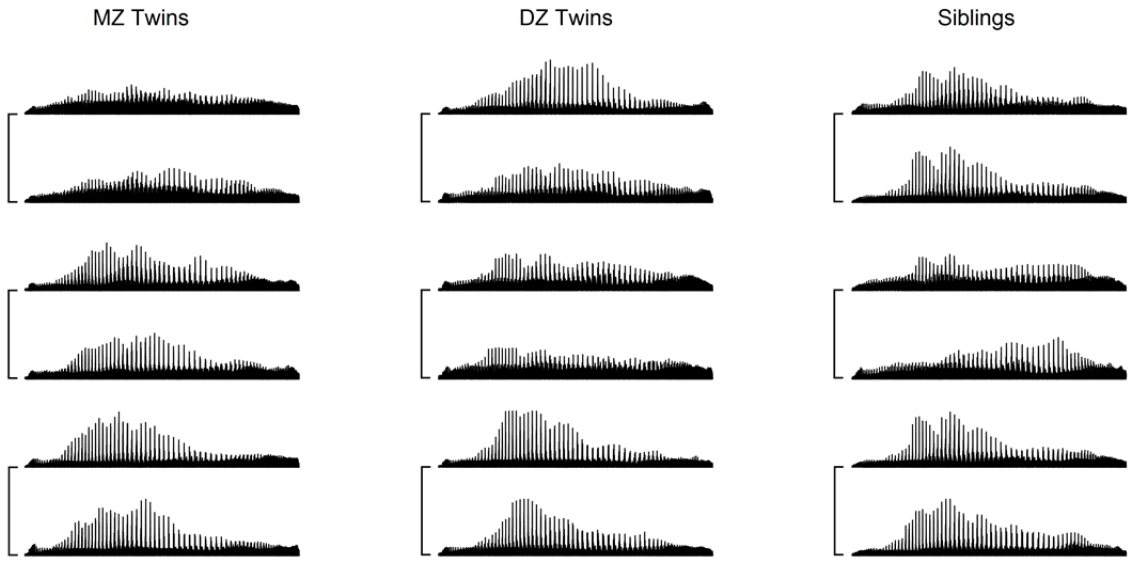

Fig 4. The local connectome fingerprints of monozygotic (MZ) twins, dizygotic (DZ) twins, and non-twin siblings.

Three pairs of connectome fingerprints are shown for each population group, and each pair is annotated by a connecting line. The connectome fingerprints between MZ twins show the grossly similar patterns though between-subject difference can still be observed. DZ twins and siblings also have a similar pattern, but the between-subject difference becomes more prominent.

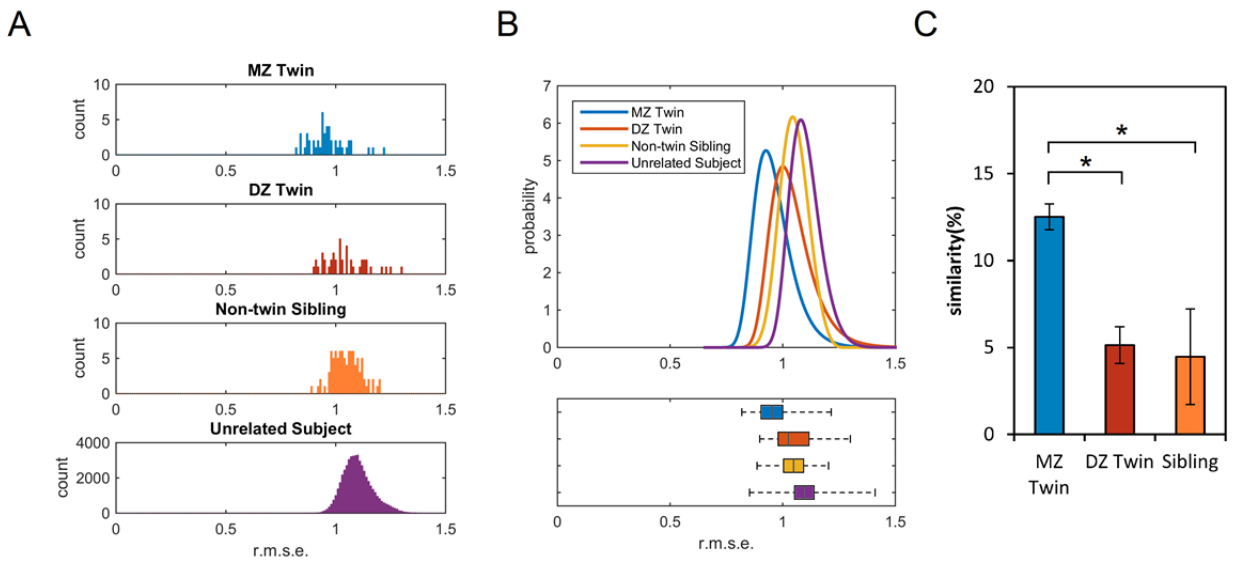

Fig 5. The similarity between genetically related subjects versus similarity between repeat scans.

(A) The histograms show the distribution of the difference between monozygotic (MZ) twins, dizygotic (DZ) twins, non-twin siblings, and genetically unrelated subjects calculated from their local connectome fingerprints. In average, MZ twins have the lowest difference between each twin pair, followed by DZ twins and siblings. (B) The upper figure shows the histograms of the difference fitted with generalized extreme value distribution, whereas lower figure shows the box plot of the distribution to facilitate comparison. The four distributions are mostly overlapped, indicating that twins and siblings still have high individuality similar to genetically-unrelated subjects. (C) The local connectome fingerprint can use as a phenotype to quantify the genetic contribution to the white matter architecture. The similarity between MZ twins is significantly higher than that between DZ twins or non-twin siblings, whereas the similarity between DZ twins is not statistically different from the similarity between non-twin siblings.

Neuroplasticity revealed by the repeated scans

Finally, we examined how time impacts the uniqueness of local connectome fingerprints. The timeframe between repeat scans varied across the four datasets. The subjects in dataset I (n=11) were scanned three times, with all sessions occurring within 16 days. The subjects in dataset II (n=24) were scanned twice with 1-3 months apart. In dataset III (n=60), 14 of the subjects were scanned twice about six months apart, and 44 of the subjects in dataset IV (n=118) were scanned

a second time one year later. Consistent with our hypothesis, the similarity in repeat scans of the same individual was strongest within 16 days (72.43% similarity in dataset I), whereas the similarity was lower at 1~3 months (69.82% dataset II), six months (67.02% dataset III) and one year (65.85% for dataset IV)(Fig. 6A). The similarity index was significantly different across all four datasets ($F[3,111] = 5.26$, $p = 0.002$). Using linear regression, we next modeled the influence of time on connectomic uniqueness (Fig. 6B). As expected, the impact of the interval on similarity index was negative (-0.01989, $p < 0.001$) meaning that within-subject similarity dropped at a rate of 0.02% per day or 7.26% per year. The longer intervals between repeat scans led to reduced similarity index of the same subject, suggesting that local connectome fingerprint may capture structural changes due to life experience.

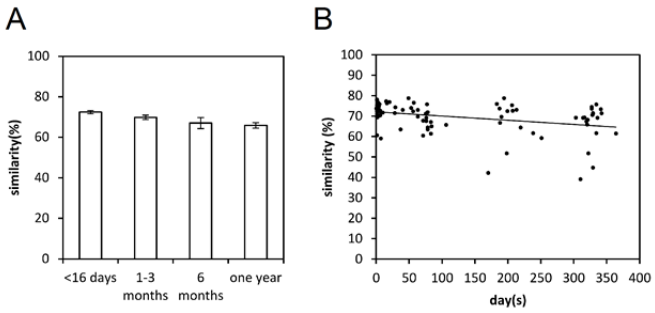

Fig 6. The similarity between repeated scans decreases at longer time intervals.

(A) In four independently acquired datasets, the similarity between repeat scan shows a decreasing trend with respect to the time interval, suggesting that local connectome fingerprint may capture neuroplasticity in the white matter architecture due to life experience. (B) The scatter plot shows the similarity of repeated scans with respect to the scan interval for each participating subjects. The longer intervals between repeat scans result in lower similarity index.

Discussion

Local white matter architecture is so unique and highly conserved within an individual that it can be considered a unique neural phenotype. Here we show that this phenotype can be quantified by measuring the density of microscopic water diffusion along major white matter fascicles and producing a high dimensional vector that can be used to compute the distance between two

structural connectomes, i.e., a local connectome fingerprint. The distance between two local connectome fingerprints reflects a low dimensional representation of both similarities and differences in whole-brain white matter pathways, Our analysis showed how the local connectome fingerprint exhibits unprecedentedly high between-subject distance, while generally low within-subject distance, allowing for it to be used as a reliable measure of the specific connective architecture of individual brains. This property paves the way for using the local connectome as a phenotypic marker of the structural connectome.

The concept of local connectome is both conceptually and methodologically different from conventional connectomic measures. While most studies have emphasized on region-to-region connectivity [3] and ignored the rich information in the local white matter architecture, the local connectome reveals the connectivity at the voxel level and characterizes local white matter architecture to provide a high dimensional data that may complement the region-to-region connectivity [11]. This local connectome mindset considers the fact that the difference between brain structures may be localized and thus may not be readily identified in the global connectomic pattern. We have previously shown that the local connectome can be used to localize the change of white matter structure due to physiological difference such as body mass index [11]. In fact, the uniqueness of fingerprints generated from the local connectome is substantially higher than what is observed in region-to-region connectivity reported by either dMRI or fMRI, as typically done in human connectomic studies [2, 16, 17]. For example, the region-to-region structural connectivity achieved a classification accuracy around 90~97%. This is very close to the accuracy of its functional counterpart [16], that was recently reported to have an accuracy of 92-94% in whole brain identification and 98-99% in frontoparietal network. Although both region-to-region connectivity approaches have accuracy greater than 90%, the performance remains substantially lower than the perfect classification in 17,398 leave-one-out rounds and an estimated error of 10^{-6} achieved by local connectome fingerprint.

At first glance, it may seem possible that the high degree of uniqueness exhibited by the local connectome fingerprint may be due to variability in the spatial normalization process between individuals driven by the unique gyral or sulcal folding patterns in gray matter. However, our comparison with the FA-based fingerprints reveals that the spatial normalization process does not fully contribute to the uniqueness observed in the local connectome fingerprint. Both FA-

based fingerprints and the local connectome fingerprints used identical spatial normalization mapping, but the FA-based fingerprints have a much higher error rate in leave-one-out cross-validation (e.g. 0.87% for dataset IV) than the zero cross-validation error achieved by the local connectome fingerprint, suggesting that a substantial portion of the uniqueness is owing to the microstructural white matter characteristics quantified in the SDF. Moreover, we get nearly as good characterization of white matter uniqueness when our analysis is restricted to a small portion of white matter with minimal influence of sulcal and gyral folding (i.e., the mid corpus callosum). In this case, the corpus callosum fingerprint itself also exhibited good uniqueness as the FA-based fingerprints sampled from the whole white matter structure. These two analyses support our claim that the local connectome fingerprint can reveal the unique characteristics of the white matter architecture.

It is important to point out that the local connectome fingerprint is sensitive to individual differences in underlying white matter architecture than more popular diffusivity-based metrics such as FA, axial diffusivity, and radial diffusivity. Diffusivity quantifies *how fast* water diffuses in tissue [18] and is sensitive to the structural integrity of the underlying fiber bundles [12], such as axonal loss and demyelination [19-22]. By contrast, SDF quantifies *how much* water diffuses along the fiber pathways [13, 23] and is sensitive to density characteristics of white matter such as the compactness of the fiber bundles [13, 24, 25]. As illustrated in our qualitative analysis (Fig. 1C), while the density characteristics vary substantially among normal populations, the diffusivity measurements do not show obvious differences between subjects. This highlights how diffusion density-based measures may be more sensitive to microstructural white matter patterns that reflect within-individual uniqueness than diffusivity-based metrics.

The high degree of uniqueness in the local connectome within an individual means that it reflects a quantifiable phenotype of neural organization. As illustrated in our analysis of twins, the similarity in monozygotic twins was around twice as much of the dizygotic twins, whereas our posthoc analysis did not find significant similarity difference between dizygotic twins and non-twins siblings. These results are highly suggestive that genetics contribute a substantial portion to the overall construction of the local connectome, which is consistent with previous studies showing high heritability in cortical connections [26, 27] and white matter integrity [28-31]. Nevertheless, our results also showed that high heritability may not necessarily imply that most

of the differences or similarity observed in phenotypes are due to genetic factors [32]. Monozygotic twins share only 12.51% similarity in local white matter architecture while repeat scans for the same subject have a much high similarity hovering around 60-70%. Thus, a considerable portion of the individuality in local connectome is likely driven by environmental factors such as life experience and learning, and monozygotic twins still exhibit high individuality in their connectome. In fact, our findings show that the local connectome fingerprint is highly plastic over time. The uniqueness of the local connectome fingerprint decreases at a rate of 0.02% per day, or 7.26% per year, raising many questions about which factors (genomic, social, environmental, or pathological) sculpt local white matter systems. Of course, white matter integrity also varies with normative development [33-35], a portion of which may be determined genetically. This warrants more longitudinal and genetic analysis to identify specific contributions of genetic and environmental factors on the uniqueness of connectomic structure, with an aim to understand how those factors interact with abnormal brain circuits in neurological and psychiatric disorders. Future work can more carefully look at how individual genetic markers contribute to the similarity in the local white matter architecture.

Methods

Five independently collected dMRI datasets

The first dataset included a total of 11 subjects (9 males and 2 females, age 20~42). Each subject had three diffusion MRI scans within 16 days on a Siemens Trio 3T system at the University of California, Santa Barbara. All methods were approved by the local institutional review board at the University of California, Santa Barbara. The diffusion MRI was acquired using a twice-refocused spin-echo EPI sequence. A 257-direction full-sphere grid sampling scheme was used. The maximum b-value was 5000 s/mm^2. TR = 9916 ms, TE = 157 ms, voxel size = 2.4×2.4×2.4 mm, FoV = 231×231 mm.

The second set of data included a total of 24 subjects (8 males and 16 females, age 22 ~ 74). All participants were scanned on a Siemens Tim Trio 3T system at National Taiwan University, and all subjects had their second scan at 1~3 months. All methods were approved by the local institutional review board at National Taiwan University. The diffusion MRI was also acquired using a twice-refocused spin-echo EPI sequence. The diffusion scheme is a 101-direction half-

sphere grid sampling scheme with b-max = 4000 s/mm^2 (b-table available at http://dsi-studio.labsolver.org). TR = 9600 ms, TE = 130 ms, voxel size = 2.5×2.5×2.5 mm.

The third set of data included a total of 60 subjects (30 males and 30 females, age 18 ~ 46). All participants were scanned on a Siemens Verio 3T system at Carnegie Mellon University, and 14 of the 60 subjects had their second scan at 6 months. All methods were approved by the local institutional review board at Carnegie Mellon University. The diffusion MRI was also acquired using a twice-refocused spin-echo EPI sequence. A 257-direction full-sphere grid sampling scheme was used. The maximum b-value was 5000 s/mm^2. TR = 9916 ms, TE = 157 ms, voxel size = 2.4×2.4×2.4 mm, FoV = 231×231 mm.

The fourth set of diffusion data included a total of 118 subjects (91 males and 27 females, age 22 ~ 55) that were also scanned on a Siemens Verio 3T system at the Carnegie Mellon University. All methods were approved by the local institutional review board at the University of Pittsburgh and Carnegie Mellon University. 44 of them had another scan after one year. The diffusion images were acquired on a Siemens Verio scanner using a 2D EPI diffusion sequence. TE=96 ms, and TR=11100 ms. A total of 50 diffusion sampling directions were acquired. The b-value was 2000 s/mm^2. The in-plane resolution was 2.4 mm. The slice thickness was 2.4 mm.

The fifth dataset was from the Human Connectome Projects (Q3, 2014) acquired by Washington University in Saint Louis and University of Minnesota. The diffusion MRI data were acquired on a Siemens 3T Skyra scanner using a 2D spin-echo single-shot multiband EPI sequence with a multi-band factor of 3 and monopolar gradient pulse. A total of 486 subjects (195 males and 291 females, age 22 ~ 36) received diffusion scans. The spatial resolution was 1.25 mm isotropic. TR=5500 ms, TE=89.50 ms. The b-values were 1000, 2000, and 3000 s/mm^2. The total number of diffusion sampling directions was 90, 90, and 90 for each of the shells in addition to 6 b0 images. The total scanning time was approximately 55 minutes. The scan data included 49 pairs of monozygotic twin, 43 pairs of dizygotic twins, and 96 pairs of siblings.

Carnegie Mellon University Institutional Review Board (IRB) reviewed the research protocol for the data analysis in accordance with 45 CFR 46 and CMU's Federal-wide Assurance. The

research protocol has been given approval as Exempt by the IRB on March 12, 2014, in accordance with 45 CFR 46.101(b)(4) (IRB Protocol Number: HS14-139).

Local connectome fingerprinting

As shown in Fig. 2A, the diffusion MRI data of each subject were reconstructed in a common stereotaxic space using q-space diffeomorphic reconstruction (QSDR)[36], a white matter based nonlinear registration approach that directly reconstructed diffusion information in a standard space:

$$\psi(\hat{\mathbf{u}}) = \left|J_{\varphi}(\mathbf{r})\right| Z_0 \sum_i \mathrm{W}_i(\varphi(\mathbf{r}))\mathrm{sinc}\left(\sigma\sqrt{6Db_i} < \hat{\mathbf{g}}_i, \frac{J_{\varphi}(\mathbf{r})\hat{\mathbf{u}}}{\left\|J_{\varphi}(\mathbf{r})\hat{\mathbf{u}}\right\|} >\right) \qquad (1)$$

$\psi(\hat{\mathbf{u}})$ is a spin distribution function (SDF)[13] in the standard space, defined as the density of diffusing spins that have displacement oriented at direction $\hat{\mathbf{u}}$. φ is a function that maps a coordinate **r** from the standard space to the subject's space, whereas J_{φ} is the Jacobian matrix of φ, and $\left|J_{\varphi}\right|$ is the Jacobian determinant. W_i is the diffusion signals acquired by a b-value of b_i with diffusion sensitization gradient oriented at $\hat{\mathbf{g}}_i$. σ is the diffusion sampling ratio controlling the displacement range of the diffusing spins sampled by the SDFs. Lower values allow for quantifying more from restricted diffusion. D is the diffusivity of free water diffusion, and Z_0 is the constant estimated by the diffusion signals of free water diffusion in the brain ventricle[36]. A σ of 1.25 was used to calculate the SDFs, and 1 mm resolution was assigned to the output resolution of the QSDR reconstruction.

A common axonal directions atlas, derived from the HCP dataset (this HCP-488 atlas is freely available at http://dsi-studio.labsolver.org), was used as a common SDF sampling framework to provide a consistent set of sampling directions $\hat{\mathbf{u}}$ to sample the magnitude of SDFs along axonal directions in the cerebral white matter. Gray matter was excluded using the ICBM-152 white matter mask (MacConnel Brain Imaging Centre, McGill University, Canada). The cerebellum was also excluded due to different slice coverage in cerebellum across subjects. Since each voxel in the cerebral white matter may have more than one axonal direction, multiple measurements

can be extracted from the SDF of the same voxel. The density measurements were sampled by the left-posterior-superior voxel order and compiled into a sequence of scalar values (Fig. 2B). Since the density measurement has arbitrary units, the local connectome fingerprint was scaled to make the variance equal to 1. The computation was conducted using DSI Studio (http://dsi-studio.labslover.org), an open-source diffusion MRI analysis tool for connectome analysis. The source code and the local connectome fingerprint data are publicly available on the same website.

Estimation of classification error

For each dMRI dataset, the root-mean-squared error between any two connectome fingerprints was calculated to obtain a matrix of paired-wise difference. The calculated difference was used as the feature to classify whether two connectome fingerprints are from the same or different person. The default linear discriminant analysis (LDA) classifier provided in MATLAB (MathWorks, Natick, MA) was used, and for each dataset, the classification error was estimated using leave-one-out cross-validation. We also used a modeling method to calculate the classification error if the leave-one-out cross- validation did not yield any classification error. The histograms of the within-subject and between-subject differences were fitted by the generalized extreme value distribution using the maximum likelihood estimator (gevfit) provided in MATLAB. To consider the non-negativity of the distribution, the estimated k parameter of the generalized extreme value distribution was set to be greater than 0. The classification error was estimated by the probability of a within-subject difference greater than a between-subject difference estimated using the generalized extreme value distribution.

Comparison with traditional connectivity matrix

To compare local connectome fingerprint with region-to-region connectivity matrix, deterministic fiber tracking[24] was applied using a 100,000 uniform white matter seeding points, a maximum turning angle of 60 degrees, and a default anisotropy threshold determined using Otsu's threshold [37]. The cortical regions were defined through a nonlinear registration between the subject anisotropy map and the HCP-488 anisotropy map in DSI Studio and parcellated using the Automated Anatomical Labeling (AAL) atlas. The matrix entries were quantified by the number of tracks ending in each of the region pairs. The root-mean-squared

error can also be calculated from any two connectivity matrices. The classification error was also estimated and compared with local connectome fingerprint.

Similarity index

The similarity index between two local connectome fingerprints was calculated by $100\% \times (1-d_1/d_0)$, where d_1 was the difference between two fingerprints, and d_0 was the expected value of the differences between unrelated subjects scanned by the same imaging protocol. The similarity between MZ twins, DZ twins, non-twin siblings, and repeated scans was calculated and compared. To further study the similarity between repeat scans, the similarity indices were regressed against their scanning time intervals to study the effect of time interval on the local connectome fingerprints.

References

1. Seung S. Connectome: How The Brain's Wiring Makes Us Who We Are: Houghton Mifflin Harcourt Publishing Company, New York; 2012.
2. 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(7):e159. Epub 2008/07/04. doi: 07-PLBI-RA-4028 [pii] 10.1371/journal.pbio.0060159. PubMed PMID: 18597554.
3. Sporns O, Tononi G, Kotter R. The human connectome: A structural description of the human brain. PLoS Comput Biol. 2005;1(4):e42. Epub 2005/10/05. doi: 10.1371/journal.pcbi.0010042. PubMed PMID: 16201007.
4. Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci. 2009;10(3):186-98. doi: 10.1038/nrn2575. PubMed PMID: 19190637.
5. Le Bihan D, Johansen-Berg H. Diffusion MRI at 25: exploring brain tissue structure and function. Neuroimage. 2012;61(2):324-41. doi: 10.1016/j.neuroimage.2011.11.006. PubMed PMID: 22120012; PubMed Central PMCID: PMCPMC3683822.
6. Jbabdi S, Sotiropoulos SN, Haber SN, Van Essen DC, Behrens TE. Measuring macroscopic brain connections in vivo. Nat Neurosci. 2015;18(11):1546-55. doi: 10.1038/nn.4134. PubMed PMID: 26505566.
7. Hagmann P, Sporns O, Madan N, Cammoun L, Pienaar R, Wedeen VJ, et al. White matter maturation reshapes structural connectivity in the late developing human brain. Proc Natl Acad Sci U S A. 2010;107(44):19067-72. Epub 2010/10/20. doi: 1009073107 [pii] 10.1073/pnas.1009073107. PubMed PMID: 20956328.
8. Sporns O. Contributions and challenges for network models in cognitive neuroscience. Nat Neurosci. 2014;17(5):652-60. Epub 2014/04/02. doi: 10.1038/nn.3690 nn.3690 [pii]. PubMed PMID: 24686784.
9. Reveley C, Seth AK, Pierpaoli C, Silva AC, Yu D, Saunders RC, et al. Superficial white matter fiber systems impede detection of long-range cortical connections in diffusion MR tractography. Proc Natl Acad Sci U S A. 2015;112(21):E2820-8. doi: 10.1073/pnas.1418198112. PubMed PMID: 25964365; PubMed Central PMCID: PMC4450402.
10. Thomas C, Ye FQ, Irfanoglu MO, Modi P, Saleem KS, Leopold DA, et al. Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited. Proc Natl Acad Sci U S A. 2014;111(46):16574-9. Epub 2014/11/05. doi: 10.1073/pnas.1405672111 1405672111 [pii]. PubMed PMID: 25368179; PubMed Central PMCID: PMC4246325.
11. Yeh FC, Badre D, Verstynen T. Connectometry: A statistical approach harnessing the analytical potential of the local connectome. Neuroimage. 2016;125:162-71. doi: 10.1016/j.neuroimage.2015.10.053. PubMed PMID: 26499808.
12. Pierpaoli C, Basser PJ. Toward a quantitative assessment of diffusion anisotropy. Magn Reson Med. 1996;36(6):893-906. Epub 1996/12/01. PubMed PMID: 8946355.
13. Yeh FC, Wedeen VJ, Tseng WY. Generalized q-sampling imaging. IEEE Trans Med Imaging. 2010;29(9):1626-35. Epub 2010/03/23. doi: 10.1109/TMI.2010.2045126. PubMed PMID: 20304721.
14. Krzanowski WJ. Principles of Multivariate Analysis: A User's Perspective: New York: Oxford University Press; 1988.
15. Jenkinson AF. The Frequency Distribution of the Annual Maximum (or Minimum) of Meteorological Elements. Quarterly Journal of the Royal Meteorological Society, . 1955;81:158-71.
16. Finn ES, Shen X, Scheinost D, Rosenberg MD, Huang J, Chun MM, et al. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat Neurosci. 2015. doi: 10.1038/nn.4135. PubMed PMID: 26457551.
17. Hagmann P, Kurant M, Gigandet X, Thiran P, Wedeen VJ, Meuli R, et al. Mapping human whole-brain structural networks with diffusion MRI. PLoS One. 2007;2(7):e597. Epub 2007/07/06. doi: 10.1371/journal.pone.0000597. PubMed PMID: 17611629.
18. Le Bihan D, Breton E, Lallemand D, Grenier P, Cabanis E, Laval-Jeantet M. MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. Radiology. 1986;161(2):401-7. Epub 1986/11/01. doi: 10.1148/radiology.161.2.3763909. PubMed PMID: 3763909.
19. Budde MD, Kim JH, Liang HF, Schmidt RE, Russell JH, Cross AH, et al. Toward accurate diagnosis of white matter pathology using diffusion tensor imaging. Magn Reson Med. 2007;57(4):688-95. Epub 2007/03/29. doi: 10.1002/mrm.21200. PubMed PMID: 17390365.

20. Sun SW, Liang HF, Trinkaus K, Cross AH, Armstrong RC, Song SK. Noninvasive detection of cuprizone induced axonal damage and demyelination in the mouse corpus callosum. Magn Reson Med. 2006;55(2):302-8. Epub 2006/01/13. doi: 10.1002/mrm.20774. PubMed PMID: 16408263.
21. Song SK, Sun SW, Ramsbottom MJ, Chang C, Russell J, Cross AH. Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. Neuroimage. 2002;17(3):1429-36. Epub 2002/11/05. doi: S105381190291267X [pii]. PubMed PMID: 12414282.
22. Song SK, Yoshino J, Le TQ, Lin SJ, Sun SW, Cross AH, et al. Demyelination increases radial diffusivity in corpus callosum of mouse brain. Neuroimage. 2005;26(1):132-40. Epub 2005/05/03. doi: S1053-8119(05)00022-4 [pii]
10.1016/j.neuroimage.2005.01.028. PubMed PMID: 15862213.
23. Callaghan PT. Principles of Nuclear Magnetic Resonance Microscopy: Oxford University Press; 1994.
24. Yeh FC, Verstynen TD, Wang Y, Fernandez-Miranda JC, Tseng WY. Deterministic diffusion fiber tracking improved by quantitative anisotropy. PLoS ONE. 2013;8(11):e80713. Epub 2013/12/19. doi: 10.1371/journal.pone.0080713
PONE-D-13-26801 [pii]. PubMed PMID: 24348913; PubMed Central PMCID: PMC3858183.
25. Yeh FC, Wedeen VJ, Tseng WY. Estimation of fiber orientation and spin density distribution by diffusion deconvolution. Neuroimage. 2011;55(3):1054-62. Epub 2011/01/15. doi: S1053-8119(11)00015-2 [pii]
10.1016/j.neuroimage.2010.11.087. PubMed PMID: 21232611.
26. Shen KK, Rose S, Fripp J, McMahon KL, de Zubicaray GI, Martin NG, et al. Investigating brain connectivity heritability in a twin study using diffusion imaging data. Neuroimage. 2014;100:628-41. doi: 10.1016/j.neuroimage.2014.06.041. PubMed PMID: 24973604; PubMed Central PMCID: PMCPMC4291188.
27. Chiang MC, Barysheva M, McMahon KL, de Zubicaray GI, Johnson K, Montgomery GW, et al. Gene network effects on brain microstructure and intellectual performance identified in 472 twins. J Neurosci. 2012;32(25):8732-45. doi: 10.1523/JNEUROSCI.5993-11.2012. PubMed PMID: 22723713; PubMed Central PMCID: PMCPMC3420968.
28. Kochunov P, Jahanshad N, Marcus D, Winkler A, Sprooten E, Nichols TE, et al. Heritability of fractional anisotropy in human white matter: a comparison of Human Connectome Project and ENIGMA-DTI data. Neuroimage. 2015;111:300-11. doi: 10.1016/j.neuroimage.2015.02.050. PubMed PMID: 25747917; PubMed Central PMCID: PMCPMC4387079.
29. Kochunov P, Glahn DC, Lancaster JL, Winkler AM, Smith S, Thompson PM, et al. Genetics of microstructure of cerebral white matter using diffusion tensor imaging. Neuroimage. 2010;53(3):1109-16. doi: 10.1016/j.neuroimage.2010.01.078. PubMed PMID: 20117221; PubMed Central PMCID: PMCPMC2888778.
30. Chiang MC, Barysheva M, Shattuck DW, Lee AD, Madsen SK, Avedissian C, et al. Genetics of brain fiber architecture and intellectual performance. J Neurosci. 2009;29(7):2212-24. doi: 10.1523/JNEUROSCI.4184-08.2009. PubMed PMID: 19228974; PubMed Central PMCID: PMCPMC2773128.
31. Sinclair B, Hansell NK, Blokland GA, Martin NG, Thompson PM, Breakspear M, et al. Heritability of the network architecture of intrinsic brain functional connectivity. Neuroimage. 2015;121:243-52. doi: 10.1016/j.neuroimage.2015.07.048. PubMed PMID: 26226088.
32. Gray JR, Thompson PM. Neurobiology of intelligence: science and ethics. Nat Rev Neurosci. 2004;5(6):471-82. doi: 10.1038/nrn1405. PubMed PMID: 15152197.
33. Craik FI, Bialystok E. Cognition through the lifespan: mechanisms of change. Trends Cogn Sci. 2006;10(3):131-8. doi: 10.1016/j.tics.2006.01.007. PubMed PMID: 16460992.
34. Tardif CL, Gauthier CJ, Steele CJ, Bazin PL, Schafer A, Schaefer A, et al. Advanced MRI techniques to improve our understanding of experience-induced neuroplasticity. Neuroimage. 2015. doi: 10.1016/j.neuroimage.2015.08.047. PubMed PMID: 26318050.
35. Simmonds DJ, Hallquist MN, Asato M, Luna B. Developmental stages and sex differences of white matter and behavioral development through adolescence: a longitudinal diffusion tensor imaging (DTI) study. Neuroimage. 2014;92:356-68. doi: 10.1016/j.neuroimage.2013.12.044. PubMed PMID: 24384150; PubMed Central PMCID: PMCPMC4301413.
36. Yeh FC, Tseng WY. NTU-90: a high angular resolution brain atlas constructed by q-space diffeomorphic reconstruction. Neuroimage. 2011;58(1):91-9. Epub 2011/06/28. doi: S1053-8119(11)00639-2 [pii]
10.1016/j.neuroimage.2011.06.021. PubMed PMID: 21704171.
37. Otsu N. A threshold selection method from gray-level histograms. IEEE Trans Sys, Man, Cyber. 1979;9(1):62-6.

Acknowledgments

The research was sponsored by the Army Research Laboratory and accomplished under Cooperative Agreement Number W911NF-10-2-0022. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. This research was supported by and NSF BIGDATA (1247658). Part of the data used this study were from the Human Connectome Project, WU-Minn Consortium (1U54MH091657). This research was supported in part by the Ruentex Group and the Ministry of Economic Affairs, Taiwan (101-EC-17-A-19-S1-175). This research was supported in part by National Institutes of Health (R01 DK095172).