Conserved and variable architecture of human white matter connectivity

Danielle S. Bassett a,⁎, Jesse A. Brown b, c, Vibhas Deshpande d, Jean M. Carlson a, Scott T. Grafton e

a Complex Systems Group, Department of Physics, University of California, Santa Barbara, CA, USA
b Center for Cognitive Neuroscience, David Geffen School of Medicine at UCLA, Semel Institute, University of California, Los Angeles, CA, USA
c Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, CA, USA
d Siemens Medical Solutions, San Francisco, CA, USA
e Department of Psychology and UCSB Brain Imaging Center, University of California, Santa Barbara, CA, USA

Abstract

Whole-brain network analysis of diffusion imaging tractography data is an important new tool for quantification of differential connectivity patterns across individuals and between groups. Here we investigate both the conservation of network architectural properties across methodological variation and the reproducibility of individual architecture across multiple scanning sessions. Diffusion spectrum imaging (DSI) and diffusion tensor imaging (DTI) data were both acquired in triplicate from a cohort of healthy young adults. Deterministic tractography was performed on each dataset and inter-regional connectivity matrices were then derived by applying each of three widely used whole-brain parcellation schemes over a range of spatial resolutions. Across acquisitions and preprocessing streams, anatomical brain networks were found to be sparsely connected, hierarchical, and assortative. They also displayed signatures of topo-physical interdependence such as Rentian scaling. Basic connectivity properties and several graph metrics consistently displayed high reproducibility and low variability in both DSI and DTI networks. The relative increased sensitivity of DSI to complex fiber configurations was evident in increased tract counts and network density compared with DTI. In combination, this pattern of results shows that network analysis of human white matter connectivity provides sensitive and temporally stable topological and physical estimates of individual cortical structure across multiple spatial scales.

© 2010 Elsevier Inc. All rights reserved.

Introduction

Recent advances in diffusion-based magnetic resonance imaging (MRI) techniques and complementary white matter tractography have made it possible to estimate the locations of anatomical highways spanning the cortex in the healthy and diseased human brain (Basser et al., 2000; Lazar et al., 2003; Behrens et al., 2003; Hagmann et al., 2003; Parker and Alexander, 2003). Collectively, these large-scale pathways form a network architecture which, far from being random or haphazard, suggests that the cortex has precise internal structure. Studying this measured architecture may therefore provide insight into how macroscopic white matter structure both constrains and facilitates healthy cognitive function (Johansen-Berg, 2009; Bandettini, 2009; Assaf and Pasternak, 2008). For example, large-scale anatomical brain networks constructed from diffusion imaging data display a clustered or community structure, where groups of brain regions are more highly connected to each other than to regions in other groups (Iturria-Medina et al., 2007; Hagmann et al., 2008; Gong et al., 2009; Zalesky et al., 2010). These clusters of highly interconnected regions are then bridged by a few important tracts, forming a topological structure which is thought to be theoretically conducive to both segregation (through clusters or modules) and integration (through connecting paths) of information processing (Sporns et al., 2000, 2004). Through a complex structure–function relationship (Honey et al., 2007; Honey and Sporns, 2008), this delicate balance of segregation and integration in anatomical architecture may form the backdrop of modular cognitive function (Simon, 1962). Despite this confluence of results, it is not yet clear how similar our measured network structure is to true cortical organization, an uncertainty which is complicated by the availability of multiple different measurement streams including diffusion tensor imaging (DTI) (Rinderer et al., 1994; Pierpaoli et al., 1996), diffusion spectrum imaging (DSI) (Wedeen et al., 2005) and a variety of high angular resolution diffusion imaging (HARDI) based methods such as Q-ball imaging, spherical deconvolution, and PAS-MRI. In the present work, we restrict ourselves to a comparison of classical DTI and the more recently developed DSI, which has been shown to better resolve diffusion directions where white matter fibers cross (Wedeen et al., 2008).

While large-scale network descriptions have enabled an understanding of the global structure of anatomical connectivity, contemporaneous studies of single tracts or small portions of white matter
have begun to elucidate the underlying forces constraining individual connectivity (Bosnell et al., 2008). In particular, it has been suggested that anatomical connectivity is not completely hard-wired for an individual’s lifetime, but instead changes appreciably with age (over long time scales) (Giorgi et al., 2010; Wozniak and Lim, 2006), during development (over intermediate time scales) (Bava et al., 2010; Cascio et al., 2007), and with rehabilitation or training (over short time scales) (Scholz et al., 2009; Bosnell et al., 2008). In addition to their sensitivity to time-dependent changes, measures of white matter integrity are also modulated by disease and cognitive ability (White et al., 2008; Sexton et al., 2009; Madden et al., 2009).

Together, these two important strands of global and local interrogation suggest that it may be possible to characterize the large-scale anatomical connectivity of an individual and to map connectivity properties to individual cognitive ability, traits, and the effects of training. Theoretically, the utility of network theory in the study of individual variation lies in its increased power to capture large-scale alterations in structure as a combined result of many small-scale changes. However, before embarking on a study of behavioural or cognitive correlates of individual network properties, it is important to assess the sensitivity of network analysis to individual variation as well as its robustness to iterative measurement. Therefore, in this study we sought to answer two distinct but complementary questions: 1) What network properties are conserved across individuals and robust to changes in image acquisition and analysis methods? and 2) Do network properties have the ability to characterize individual differences that are both accurate and reproducible in large-scale cortical structure across multiple scanning sessions?

In order to address these questions, we acquired diffusion imaging scans from healthy individuals and systematically varied our processing stream to ascertain the effects of imaging acquisition (DSI/DTI), atlas, and spatial resolution on measured cortical architecture over the whole population as well as on its reproducibility in a single individual. More specifically, both diffusion tensor imaging (DTI) and diffusion spectrum imaging (DSI) scans were each acquired in triplicate from a group of healthy young adult individuals. Deterministic tractography was used to construct subject-specific networks and a whole-brain atlas was then applied to the tractography data in order to attain an inter-regional connectivity matrix; see Fig. 1 for a schematic of brain network construction. Atlases used in this work included the Automated Anatomical Labeling Atlas (AAL), the Harvard–Oxford Atlas (HO) and LONI Probabilistic Brain Atlas (LPBA40). In order to probe the effects of spatial resolution, each atlas was iteratively upscaled into twice as many regions of interest (ROIs) (‘Sub 1’), four times as many ROIs (‘Sub 2’), and eight times as many ROIs (‘Sub 3’); see Fig. 1 for a schematic of the upsampling scheme and Materials and methods for details.

Structural properties of cortical connectivity were measured on both raw connectivity matrices and binarized brain graphs. We hypothesized that while exact values of network metrics might vary from individual to individual, qualitative architectural properties would be conserved across the population and robust to effects of individual variability, diffusion scanning technique, and methodological variation. To complement topological analysis, we also tested for the existence of relationships between graph measures of connectivity and physical measures of connectivity, such as the length of connections in stereotactic space. Finally, in order to determine whether network properties can accurately measure individual differences, we used the intra-class correlation coefficient (ICC) to quantify the reproducibility of network measures across the 3 scanning sessions and its dependence on type of acquisitions (DTI and DSI), anatomical parcellation and spatial resolution. Specifically, by using the ICC we were able to test whether the variance between subjects in a given network property was larger than the variability within a subject over scanning sessions. By these means both the conserved and individually varying properties of healthy brain architecture, robust to imaging modality, atlas, and spatial scale, could be identified.

### Materials and methods

#### Data acquisition

After having obtained informed consent, seven healthy volunteers each completed three DSI scans acquired on separate days (mean time from first scan to last was 21 days, range 11–37). In addition, 6 of the 7 subjects also completed three DTI scans acquired on separate days (mean time from first scan to last was 5.5 days, range 3–9). All scans were acquired at 3 T with a Siemens Tim Trio MRI scanner with a 12 channel phased array head coil using an echo-planar diffusion-weighted technique acquired with iPAT and an acceleration factor of 2.

For the DSI, the timing parameters of the pulse sequence were TE/TR = 133/10,500 ms, 258 diffusion directions with a maximal b-value of 5000 s/mm²; for the DTI, the timing parameters of the pulse sequence were TE/TR = 94/8400 ms, 30 diffusion directions with a maximal b-value of 1000 s/mm² and two averages. A single b0 image was acquired during the DSI scan, and two were acquired during the DTI scan. For both scans, the matrix size was 128 × 128 and the slice number was 60. The field of view was 230 × 230 mm² and the slice thickness 2 mm. Acquisition time: 45:42 min per DSI scan and 9:08 min per DTI scan. See Supplementary Materials for a plot of b-values used in both DSI and DTI acquisitions; note that points were taken on a Cartesian grid.

In addition to diffusion scans, a three dimensional (3D) high-resolution T1-weighted sagittal sequence image of the whole brain was obtained by a magnetization prepared rapid acquisition gradient-echo (MPRAGE) sequence with the following parameters: TR = 15.0 ms; TE = 4.2 ms; flip angle = 9°, 3D acquisition, FOV = 256 mm; slice thickness = 0.89 mm, matrix = 256 × 256.

#### Data preprocessing

Motion artifact and image distortions caused by eddy-currents were corrected in FMRIB’s Diffusion Toolbox in FSL software by applying an affine alignment of each diffusion-weighted image to the b0 image. In the current study, we did not correct for EPI distortions. In this Siemens scanner, the geometric distortion for diffusion imaging from EPI was found in prior tests to be less than 2 mm (i.e., less than a single voxel), and mainly along the anterior posterior (phase encoding) direction. Because the resolution of the diffusion images was larger than the magnitude of the distortion, no correction was required.

#### Tractography

Reconstruction of the diffusion images was performed using Diffusion Toolkit (DTK) (Ruopeng Wang, Van J. Wedeen, TrackVis.org, Martinos Center for Biomedical Imaging, Massachusetts General...
Hospital (MGH) (Wang et al., 2007), a recently constructed software toolbox which provides precise diffusion imaging analysis and visualization capabilities (Granziera et al., 2009; Wedeen et al., 2008; Wahl et al., 2009; Vishwas et al., 2010; Lagana et al., 2010; Nezamzadeh et al., 2010). Diffusion tensor estimation was performed using the linear least-squares fitting method (Wang et al., 2007). The raw data was not smoothed or sharpened prior to reconstruction. Deterministic tractography was subsequently performed in TrackVis software using the fiber assignment by continuous tracking (FACT) algorithm (Mori et al., 1999; Mori and van Zijl, 2002; Xue et al., 1999). In this process, a single seed was placed in the center of each voxel, and the path was continued in the adjacent voxel which minimized the path curvature; paths were terminated for curvatures greater than 35°. Fiber tracts which were rejected by the algorithm, for example due to high curvature, were not included in the present analysis. Similarly, fiber tracking for DSI data was performed using a FACT-like algorithm in which each voxel was treated as having multiple principle diffusion directions (Wang et al., 2007), and a path was seeded for every orientation density function (ODF) maximum vector in each voxel. No further fine tuning of parameters within the software toolbox was performed. Importantly, in using DTK, we performed an exhaustive search approach in which fiber tracking is performed in all voxels, rather than in a priori specified regions of interest. Inter-regional connectivity was then examined by applying a set of gray matter masks to the complete tractography solution and counting the number of tracts which passed between any two masks.

Inter-regional connectivity

In order to attain regional rather than voxel-based connectivity, a set of N brain region masks were applied to the reconstructed fiber tracts, using the UCLA Multimodal Connectivity Package (UMCP). We determined the number of tracks that originate in one mask, i, and terminate in another mask, j, for all possible pairs of N masks, creating an N × N inter-regional anatomical connectivity matrix, Mij, where the value of any element of the matrix Mij is equal to the number of tracts originating in mask i and terminating in mask j. The matrix sum of the connectivity matrix is the total number of reconstructed fiber tracts, F, between gray matter regions uncovered by the algorithm was data driven rather than defined a priori, and was therefore variable from individual to individual and from scan to scan. In addition to the connectivity matrix, we determined the tract length matrix, Tij, where the value of any element of the matrix Tij is equal to the average length of the tracts originating in mask i and terminating in mask j.
Brain region masks

Masks for each of the $N$ regions of a cortical atlas were transformed into the subject’s native space in a multi-step process (Gong et al., 2009; Shu et al., 2009). The subject’s MPRAGE scans from the three sessions were averaged together to create a mean structural scan, which was then registered to the subject’s b0 diffusion image using the affine transform provided by FSL’s linear registration toolbox, FLIRT. This native space MPRAGE was then registered to the nonlinear Montreal Neurological Institute 152 T1 2 mm brain. The inverse of this transformation matrix was then used to warp the atlas ROIs into subject native space.

Atlasses and upsampling procedure

Three separate atlases were used in this study: 1) the Automated Anatomical Labeling Atlas (AAL) (Tzourio-Mazoyer et al., 2002; Desikan et al., 2006), 2) the Harvard–Oxford Atlas (HO), and 3) the LONI Probabilistic Brain Atlas (LPBA40) (Shattuck et al., 2007). The AAL atlas included 90 cortical and subcortical regions, the HO atlas 110 cortical and subcortical regions, and the LPBA40 54 cortical regions; none of the atlases included cerebellar structures or the brainstem. In order to assess the effect of spatial resolution, we further upsamplded each atlas to 3 additional levels of spatial resolution: 1) Sub 1, where each region of the original atlas is cut into 2 equally sized regions, 2) Sub 2, where each region of the Sub 1 atlas is cut into 2 equally sized regions, and 3) Sub 3, where each region of the Sub 2 atlas is cut into 2 equally sized regions. For example, the AAL (N=90) atlas was decomposed into AAL Sub 1 (N=180), AAL Sub 2 (N=360), and AAL Sub 3 (N=720), and similarly for the other two atlases (for HO, N=110, 220, 440, and 880 and for LPBA40, N=54, 108, 216, and 432). The upsampling algorithm used took a single brain region, found its principal spatial axis, and for LPBA40, N=54, 108, 216, and 432). The upsampling algorithm similarly for the other two atlases (for HO, N=110, 220, 440, and 880 and for LPBA40, N=54, 108, 216, and 432).

Network properties

Basic connectivity properties

Before turning to complex graph metrics, we can quantify several basic connectivity properties of both the connectivity matrix, $M_{ij}$, including the density, average weight, nodal diversity, and edge diversity and the tract length matrix, $T_{ij}$.

Density. Arguably one of the most simplistic measures of network structure is its density, simply defined as the number of non-zero edges in the graph, $E$, divided by the total number of possible edges in the graph:
\[
\kappa = \frac{E}{N(N-1)},
\]

where $N$ is the number of nodes in the graph, or in this case the number of or brain regions (Coleman and More, 1983). The density is therefore proportional to the total number of connected pairs of brain regions, irrespective of the number of tracts passing between those pairs.

Weight. The average weight is a normalized version of the total weight (Garlaschelli, 2009), and is given by
\[
\overline{W} = \frac{1}{N(N-1)} \sum M_{ij}
\]

where $i$ and $j = 1, 2, 3, \ldots N$ and $M_{ij}$ is the connectivity matrix. The average weight is therefore proportional to the total number of reconstructed white matter tracts, $F$. A measure of the number of reconstructed white matter tracts per brain region, the nodal strength of node $i$ is given by the column sum of the connectivity matrix:
\[
S(i) = \frac{1}{N-1} \sum M_{ij}
\]

Diversity. Diversity measures were originally introduced in the context of social networks (Campbell et al., 1986), and have more recently been applied to the study of functional brain networks in schizophrenia (Lynall et al., 2010). The nodal diversity is defined as the variance, over nodes, of the nodal strength
\[
D_n = \left[ \frac{1}{N} \sum (S(i) - \bar{S})^2 \right]^{1/2},
\]

while the edge diversity is given by the column variance of $M$:
\[
D_e(i) = \left[ \frac{1}{N} \sum (M_{ij} - S(i))^2 \right]^{1/2}.
\]

Tract length. The average tract length is a normalized measure of the total tract length, and is defined as
\[
T = \frac{1}{N(N-1)} \sum T_{ij},
\]

where $i$ and $j = 1, 2, 3, \ldots N$ and $T_{ij}$ is the tract length matrix, where the value of any element of the matrix $T_{ij}$ is equal to the average length of the tracts originating in mask $i$ and terminating in mask $j$.

Binary graph properties

Graph metric properties were determined in two ways: as single values on unthresholded binary graphs and as a range of values on thresholded binary graphs. Unthresholded binary graphs are perhaps the simplest to construct, and are formed such that two regions are considered to be connected with a strength of unity if at least one reconstructed tract was found linking them. These simple graphs provide a direct representation of the full complex topological organization of the anatomical brain network and are useful for determining basic structural characteristics such as small-worldness, assortativity, or hierarchy.

Unthresholded networks were mathematically defined as those in which a binary edge was said to exist if at least 1 tract was found connecting region $i$ to region $j$, and the density of these networks was denoted by $\kappa = \kappa_{\text{max}}$. In the current work, the average $\kappa_{\text{max}}$ was consistently less than the theoretical maximum of 1, which would indicate a completely connected graph. This property of anatomical matrices derived from tractography, of having a maximum density less than the theoretically possible maximum density of 1, is unlike association matrices derived from functional data. For example, an association matrix defined by the temporal correlation between regional time series provides a completely connected network where all possible connections between any two regions, $i$ and $j$, are given a weight equal to the correlation between the time course of region $i$ and the time course of region $j$ (Achard et al., 2006). For anatomical networks derived from tractography, however, many connectivity matrix elements, $M_{ij}$, are equal to zero because no tract was found connecting region $i$ to region $j$.

However, unthresholded binary graphs do not retain any information regarding the weight of any of the edges which make up the network. For this reason, we also studied thresholded binary graphs which are constructed over a range of connection densities by iteratively thresholding the weight matrix to retain connections with stronger or weaker edge weights in the resultant network.

Mathematically, in order to assess the network structure over a range of network densities, the connectivity matrix was thresholded to retain a given percent of strongest connections; the percent of connections remaining (also known as the density, or network cost $K$) ranged from 0 to a maximum comparable density $K_{\text{com}}$ defined as the maximum density common to all subjects in the particular atlas and resolution under study. The range $0 < K < K_{\text{com}}$ was defined as the
available cost regime. It is important to note that in order to achieve a constant $K$ across subjects, different thresholds were applied to different subjects. The requirement of a common density across subjects in a given atlas and resolution ensured that subsequently determined network metrics were mathematically comparable across subjects. This requirement was particularly necessary to ensure that the reproducibility of graph metrics over scanning sessions was not purely predicted by changes in density.

Having constructed either unthresholded or thresholded binary graphs, we next examined the organization of these graphs using a variety of graph metrics recently proposed in the literature (Rubinov and Sporns, 2009; Deuker et al., 2009): pathlength and clustering coefficient (Watts and Strogatz, 1998), global and local efficiency (Latora and Marchiori, 2001; Achard and Bullmore, 2007; Bassett et al., 2009), betweenness centrality (Freeman, 1977), modularity (Leicht and Newman, 2008; Meunier et al., 2009b), hierarchy (Ravasz and Barabasi, 2003; Bassett et al., 2008), synchronizability (Barahona and Pecora, 2002; Bassett et al., 2006), assortativity (Newman, 2006; Bassett et al., 2008), and robustness (Achard et al., 2006; Lynall et al., 2010). In addition to topological metrics, we also studied the purely physical measure of mean connection distance (Bassett et al., 2008), as well as a combined topo-physical property known as Rent’s exponent (Bassett et al., 2010) which provides information regarding the efficiency of topological embedding into a physical space. Because the majority of these metrics have been previously defined and widely used, we will only define those which are of particular importance to the interpretations and conclusions presented in this paper, including hierarchy, assortativity, mean connection distance, and Rent’s exponent. See Supplementary Materials for detailed mathematical definitions of all other properties.

Degree. To begin, a network is composed of units, called nodes, and connections between those units, called edges. The number of edges emanating from a node, $i$, is called the degree of node $i$, and is commonly denoted by $k_i$.

Pathlength. The pathlength between nodes $i$ and $j$ is defined as the shortest number of edges one would have to traverse in order to move along the network from node $i$ to node $j$ (Dijkstra, 1959). In binary networks, the pathlength, $L_{ij}$, between any two nodes is therefore an integer and the pathlength of an entire network, $L$, is defined as the average pathlength from any node to any other node in the network:

$$L = \frac{1}{N(N-1)} \sum_{i \neq j} L_{ij}.$$

Clustering coefficient. The clustering coefficient, $C$, is defined as the ratio of the number of connected triples to the number of triangles (Watts and Strogatz, 1998). A connected triple is a set of 3 nodes, the pathlength between any of which is less than infinity (for example, imagine nodes 1, 2, and 3 in a line with node 1 connected to node 2 and node 2 connected to node 3). A triangle is a particular type of connected triple where each node is connected to both of its neighbors, the trio forming a fully connected triangle.

Hierarchy. Networks may display a hierarchical structure in which hubs extend many long-distance connections but lack a degree of local connectivity while non-hubs extend only a few connections largely to their immediate neighbors. This hierarchical structure of a network can be characterized by the coefficient $\beta$, which is a parameter quantifying the power-law relationship between the clustering coefficient $C$ and the degree $k$ of all nodes in the network (Ravasz and Barabasi, 2003):

$$C \sim k^{-\beta}.$$  \hspace{1cm} (7)

Assortativity. The degree assortativity of a network, or more commonly simply ‘assortativity’, is defined as

$$r = \frac{\sum_{i,j} E^{-1} L_{ij} (k_i - \bar{k})(k_j - \bar{k})}{\sum_{i,j} E^{-1} L_{ij} (\bar{k})^2},$$  \hspace{1cm} (8)

where $k_i, k_j$ are the degrees of the nodes at either end of the $e$th edge, with $i = 1...E$ (Newman, 2006). The assortativity measures the preference of a node to connect to other nodes of similar degree (leading to an assortative network, $r > 0$) or to other nodes of very different degree (leading to a disassortative network, $r < 0$). Social networks are commonly found to be assortative while networks such as the internet, World-Wide Web, protein interaction networks, food webs, and the neural network of Caenorhabditis elegans are disassortative.

Mean connection distance. The estimated connection distance of an edge, $d_{ij}$, is defined as the Euclidean distance between the centroids of the connected regions $i$ and $j$ in standard stereotactic space. The mean connection distance, $d$, is defined as the average connection distance over all edges in a network (Bassett et al., 2008). Thus connection distance differs from the other, topological and dimensionless graph metrics in that it represents a spatial or topographic property of the network and has units of distance (mm) or voxels. The mean connection distance also differs from the average tract length, $T$, in that it provides a standardized estimate of the physical extent of network structure which is comparable across subjects while the tract length provides a direct measurement of connection lengths in an individual subject’s brain. While the two measures therefore provide complementary theoretical value, they are in practice highly correlated (see Supplementary Fig. 2).

Rent’s exponent. Rent’s exponent is a topo-physical property of a network; that is, it describes how a non-physical topology is embedded into a physical space, which in the case of neuronal fiber tracts is the physical space of the brain (Bassett et al., 2010). Rent’s rule, first discovered in relation to computer chip design, defines the scaling relationship between the number of external signal connections (edges), $e$, to a block of logic and the number of connected nodes, $n$, in the block (Christie and Stroobandt, 2000):

$$e \sim n^p,$$  \hspace{1cm} (9)

where $0 \leq p \leq 1$ is the Rent exponent. Following Bassett et al. (2010), the Rent’s exponent is found by tiling the Euclidean space of the network with $N_{box}=5000$ overlapping randomly sized boxes (e.g., three-dimension cubes). In each box we determine the number of nodes ($n$) and the number of connections ($e$) that cross the box boundaries. The gradient of a straight line fitted to $\log(n)$ vs $\log(e)$ using iteratively weighted least-squares regression is an estimate of the Rent exponent $p$. To minimize boundary effects, $p$ is estimated using the subset of boxes which contains less than half the total number of nodes, $n < N/2$.

Random network null models

In order to determine whether the organization structure apparent in brain networks was significant, we constructed two types of random network null models: pure random networks and pseudo random networks. Pure random networks were composed of an identical number of undirected (symmetric) edges as the true brain networks, but those edges were randomly distributed across all possible pairs of nodes. Pseudo random networks were constructed to retain not only the density of the true brain network but also the degree distribution. We used the process proposed by Maslov and Sneppen (2002) to construct pseudo random networks by rewiring
individual edges by careful substitution in order to preserve the true degree distribution. In our calculations, each edge in the real network was thus rewired at least 20 times, ensuring significant random structure within the confines of the empirical degree distribution.

**Measures of individual variability**

In the present study, we used several measures to characterize the individual variability of connectivity patterns or graph metrics including a general measure of matrix similarity and two statistical measures of reproducibility: the intra-class correlation coefficient and the coefficient of variation. In an initial analysis of the similarity between connectivity matrices themselves, we computed the (Pearson’s) correlation coefficient between the edge weights of two different matrices, either from different scanning sessions to determine inter-scan variability or from different subjects to determine inter-subject variability. This method has previously been used to show a higher similarity between scanning sessions in a single individual than between the same scanning session in different individuals (see Supplementary Information Fig. S2 in Hagmann et al. (2008)) in a DSI network analysis study.

In order to specifically test the hypothesis that variation within an individual between scanning sessions is larger than the variation across individuals, we used the statistical measure of the intra-class individual between scanning sessions is larger than the variation accounted for by between-subject variation (Lachin, 2004; McGraw and Wong, 1996). The ICC is defined as:

$$\text{ICC} = \frac{\sigma_{bs}^2}{\sigma_{bs}^2 + \sigma_{ws}^2}$$

where $\sigma_{bs}$ is the between-subject variance and $\sigma_{ws}$ is the (pooled) within subject variance. While the ICC is a normalized measure and therefore has a maximum of 1, values above 0.5 indicate that there is more variability between subjects than between scans.

Throughout this work, we determined the reproducibility of basic connectivity and graph metric values over scanning sessions, and these metrics can be placed into two main categories: global network measures, where a single network metric value is attained per network, and regional measures, where a value of a metric is obtained for each brain region in the atlas. Global measures which were tested for reproducibility include the density, average weight, nodal diversity, edge diversity, pathlength, global efficiency, betweenness, clustering, local efficiency, modularity, hierarchy, synchronizability, assortativity, robustness, Rent’s exponent, and mean connection distance. The sole regional measure tested for reproducibility was nodal strength.

A complementary statistic, the coefficient of variation (CV) is a normalized measure of the variability of a metric. The CV indicates the minimum percentage signal change detectable in repeated measures and is defined as the mean within subject standard deviation, $\sigma_{ws}$, divided by the overall measurement mean, $\mu$. (Lachin, 2004):

$$\text{CV} = \frac{\sigma_{ws}}{\mu}$$

As a result, the CV is particularly useful in comparing results between datasets with widely different means.

**Statistics and software**

All computational and simple statistical operations (such as t-tests and correlations) were implemented using MATLAB (2007a, The MathWorks Inc., Natick, MA) software. Graph metrics were estimated using a combination of in-house software, the Brain Connectivity Toolbox (Rubinov and Sporns, 2009), and the MATLAB Boost Graph Library (http://www.stanford.edu/~dgleich/programs/matlab_bgl/). The repeated measures ANOVA was performed using Statistica (version 9, StatSoft Inc.) software.

**Visualization**

Brain networks were visualized using both ball-and-stick methods (gplot.m, MATLAB File Exchange http://www.mathworks.com/matlabcentral/fileexchange/6793-gplot) and surface projections (Caret, Van Essen Laboratory, http://brainvis.wustl.edu/wiki). The mapping of the image volume to the surface was performed using the Caret PALS-B12 (Population-Average, Landmark- and Surface-based) atlas, which is derived from 12 healthy young adult subjects. The image volume is mapped to each subject separately, and the average surface rendering is used for visualization; the inter-region boundaries are therefore blurred due to inter-subject differences, and the complete mapping therefore gives a realistic depiction of group effects.

**Results**

**Conserved connectivity structure**

Graph analysis is a powerful data reduction technique, and the wide variety of available metrics provides a compelling means of characterizing connectivity profiles of both anatomical and functional neuroimaging data. However, no single metric provides tight constraints on the interpretation of the true architecture of the cortex; instead, a diverse ensemble of architectures is consistent with a given value of any particular metric. For example, two systems with equal values of global efficiency could produce that value by very different architectural means, i.e., through very different topologies. In combination, however, graph metrics can provide meaningful constraints on our understanding of the systems architecture of the cortex. In order to uncover which constraints on anatomical structure are measured by diffusion imaging, we first quantified the most simplistic characteristics of cortical architecture and progressively built a more detailed understanding using more complicated measures.

We begin our foray into the architecture of human brain anatomical networks by considering unthresholded binary graphs, where two regions are considered to be connected with a strength of unity if any reconstructed tracts were found linking them. Such a definition is arguably the simplest path to anatomical brain network construction and in subsequent sections we will consider more nuanced network definitions.

**Conserved topological architecture**

Based on a simple measure of the number of inter-regional connections, all networks studied, across individuals, diffusion scanning techniques, atlases, and resolutions, were sparse: that is, brain regions or nodes were usually connected to fewer other brain regions. More specifically, the density of a network, or the proportion of connected pairs of brain regions, was relatively low with an average value of 0.15; see Fig. 2A and Table 1. Sparse networks, unlike fully connected networks, may vary topologically from perfectly random to highly organized. It is therefore also important to note that all brain networks constructed in this study showed hallmarks of non-random topology (Fig. 2B) which have been consistently identified in similar networks previously, such as a normalized clustering coefficient $(C/C_r)$ greater than one and a normalized pathlength $(L/L_r)$ roughly equal to one (Hagmann et al., 2008; Gong et al., 2009; Iturria-Medina et al., 2007). However, this combination of high clustering and short pathlength, also known as ‘small-worldness’, is found in systems with very different inherent network architectures, and we therefore next turned to a characterization of the local connectivity of individual nodes in the anatomical network.

Topological hierarchy indicates a structured local connectivity where ‘provincial’, or less connected, nodes have disproportionately high clustering (Ravasz and Barabasi, 2003); see Fig. 2C, top. This
property, evident in all networks studied here (see Fig. 2C, bottom), is more specifically defined as a *negative* relationship between the degree of a node (one mathematical predictor of its pathlength) and its clustering coefficient, and thereby combines the two global properties we just studied. To put this result in context, it is important to note that pure random networks do not display any significant relationship between clustering and degree. For the majority of density values, the slope of a line fitted through the clustering vs degree plot should remain close to zero, although for very sparse random networks, a non-significant small negative slope may exist due to finite size effects. In addition to hierarchical and random networks, alternative network architectures may display inverse hierarchies, characterized by a positive relationship between clustering and degree (see Bassett et al. (2008) for a clinical example).

While the hierarchical architecture quantifies the nodal building blocks of the network, assortative mixing provides information about what types of nodes connect to one another (Newman, 2002). The simplest type of assortativity is called *degree assortativity*, which quantifies the relationship between a node’s degree and the average degree of its neighbors. All networks in this study consistently demonstrated an assortative nature (see Fig. 2D, bottom), indicated by a positive correlation between these two topological metrics. In contrast, pure random networks do not display a correlation between nodal degree and neighbor degree, and therefore assortativity values remain close to zero throughout the full density range.

In synthesizing this sequence of results, we find that we are studying a sparse, highly clustered, hierarchical, assortative network, for which we can construct a hypothetical schematic, as given in Fig. 2D, top. Importantly, the combination of these properties is consistent with the hypothesis that anatomical cortical networks are hierarchically modular systems: the whole system is made up of small modules or subsystems, which are in turn made up of even smaller modules or subsystems.
Conserved topo-physical relationships

The network properties considered thus far do not give us any information about how this complex network topology is mapped within the physical boundaries of the cortex. Previous work in anatomical neural networks indicates that long-distance connections in both topological (Watts and Strogatz, 1998; Hilgetag et al., 2000) and physical space (Sakata et al., 2005; Kaiser and Hilgetag, 2006; Chen et al., 2006; He et al, 2007; Bassett et al., 2008) are rare, arguably due to energetic constraints (Bassett et al., 2009; Attwell and Laughlin, 2001; Niven and Laughlin, 2008). Interestingly, across atlases, resolutions, and modalities, topological and physical distance were inversely correlated in the anatomical networks in this study (Fig. 3A): a ‘provincial’ node with predominantly short range physical connections was topologically more separated from other nodes by a large pathlength, than nodes with longer range physical connections.

It has been previously suggested that the hierarchically modular structure of the human brain is efficiently mapped into physical space (Bassett et al., 2010), which suggests there may be further relationships between local connectivity, or clustering, and physical distance. A negative correlation between these two metrics would indicate that highly clustered nodes have neighbors close by in physical space, suggesting that anatomical modules are spatially localized within the cortex. Indeed, networks constructed in this study displayed a significant negative correlation between nodal clustering and nodal connection distance which was increasingly evident when using finer spatial parcellation schemes (see Fig. 3B). Importantly, the relationships between clustering or pathlength and connection distance, as estimated by the Euclidean distance between the centers of regions, were further confirmed by significant correlations between these topological metrics and average tract length as estimated during tractography (see Supplementary Material).

These simple relationships suggest a very strong link between topology and physical space in the human brain. Interestingly, such a link has also been described in very large-scale integrated circuits where it is commonly tested for by assessing the presence or absence of Rentian scaling. Rent’s rule describes a relationship between the number of nodes in a physical patch of the network, and the number of edges crossing the boundary of that patch (Bassett et al., 2010); see Fig. 3C for a schematic. The presence of Rentian scaling in a system suggests that the network topology has been mapped cost-efficiently into physical space. Indeed, this measure of topo-physical interdependence was evident in both DSI and DTI networks across all atlases and resolutions (see Fig. 3D).

Modulation of connectivity structure

While our results suggest that many architectural principles of anatomical networks are highly conserved across modalities, atlases, and spatial resolutions, specific network metric values can be modulated by changes in network size (e.g., number of nodes and edges) even if the underlying topology of the graph remains constant (Anderson et al., 1999). In this study, network size changed appreciably over spatial resolutions: atlases with higher spatial resolution contained more nodes than atlases with lower spatial resolution. As might therefore be expected, exact values of the majority of graph metrics used in this study were strongly modulated by spatial resolution (see, for example, the normalized clustering coefficient, or clustering). Indeed, networks constructed with the HO atlas applied to DSI (red) and DTI (black) data (see Supplementary Figs. 4–7 for comparable results from other atlases and resolutions in addition to using estimated tract length rather than Euclidean distance). (A) Relationship between the topological and estimated physical length (left: DSI, \( r = -0.45, p = 4 \times 10^{-7}; \) DTI, \( r = -0.69, p = 2 \times 10^{-11} \)) of a node’s connections. (B) Relationship between the clustering coefficient and estimated physical length (left: DSI, \( r = -0.47, p = 1 \times 10^{-7}; \) DTI, \( r = -0.61, p = 5 \times 10^{-13} \)) of a node’s connections. (C) In order to determine the presence or absence of Rentian scaling, the physical brain network is partitioned into randomly sized cubes (top). The number of nodes, \( n \), inside each cube and the number of edges, \( e \), crossing the boundary of each cube are counted (bottom). (D) Rentian scaling is then defined as a scaling relationship between the number of edges (\( e \)) crossing the boundary of the arbitrarily sized region of cortex and the number of nodes (\( n \)) within that region: \( e = n^r \). Here we show the results for the HO original network of a single subject. Note that Rentian scaling did not seem to be largely modulated by imaging modality; the average Rent’s exponent in DSI networks was \( p = 0.79 \pm 0.03 \) and for DTI networks was \( p = 0.81 \pm 0.03 \) (3D over subjects). See Supplementary Fig. 8 for comparable results from other atlases and resolutions.
The relationship between graph metric values and the number of nodes in the atlas is, however, complicated by the inextricable link between nodes and edges: atlases with higher spatial resolutions not only contained more nodes but also contained a lower density of edges. In fact, the number of nodes in an atlas was inversely correlated with the connectivity density: Pearson’s correlation coefficient over atlases, resolutions, and modalities was $R = -0.69$, $P = 0.0002$. Upon inspection, the relationship between the density and the number of nodes was found to follow a power-law with similar exponents in both DSI and DTI networks; see Supplementary Fig. 3. By extension, network metrics highly dependent on density also displayed power-law scaling with the number nodes in the atlas. Due to these complicated interactions between graph metric values and the variable size of the networks under study, we were unable to make further comparisons between network property values across atlases.

**Variability and reproducibility of raw connectivity**

In a second branch of analysis, we were specifically interested in determining whether connectivity properties which were conserved across the population could be reproducibly quantified in an individual over multiple scanning sessions, and whether that reproducibility could be potentially influenced by methodological variation. We began by assessing the variability and reproducibility of the raw connectivity matrices before turning to binary graph metrics.

**Connectivity matrices**

Before addressing complex network architecture, we focused on the raw connectivity matrices themselves where we noted that connectivity matrices derived from DSI data had consistently higher weights than those derived from DTI data, indicating a relative increase in the number of reconstructed tracts; see Fig. 4A. In fact, the average weight of DSI connectivity matrices was $\sim 3.5 \times$ higher than DTI matrices (one-sample t-test, $t = 2.23$, $P = 0.047$), consistent with both the increased number of diffusion directions and ability to

---

**Fig. 4.** Individual variability in measured cortical architecture. (A) Sample connectivity matrices reconstructed for the same subject in two different sessions using DSI data (top row) and for two different subjects in the same session using DTI data (bottom row). (B) Sample correlation between the edge weights in connectivity matrices from the same subject in different scanning sessions (left) and different subjects in the same scanning session (right). (C) Boxplots indicating the average similarity between connectivity matrices derived from the same subjects (‘Within’) and derived from different subjects (‘Between’) for DSI (left) and DTI (right) data for all atlases and resolutions. Plots of the average value (D), reproducibility (E), and variability (F), of the basic connectivity properties for DSI (red) and DTI (black) networks. Error bars indicate standard deviations in the mean over atlases and resolutions.
resolve complex fiber configurations. In addition to track numbers, the density of a network derived from DSI was consistently higher than that of a network derived from DTI (one-sample t-test, \( t = 3.19, P = 0.0085 \)), as was the nodal diversity (\( 3.0 \times \) higher; \( t = 5.40, P = 2e^{-4} \)), and edge diversity (\( 3.6 \times \) higher; \( t = 2.71, P = 0.02 \)).

In addition to the differences in connectivity between the two modalities, we also noted that networks derived from the same subject were more similar than those derived from different subjects, suggesting that the pattern of connectivity contained in these matrices is highly specific to a given individual; see Fig. 4A. We quantified this similarity between any two connectivity matrices by computing the correlation coefficient between their respective weights (Hagmann et al., 2008); see for example Fig. 4B and Materials and methods for details. As might be expected, the average similarity between connectivity matrices derived from the same subject was consistently higher than the similarity between matrices derived from different subjects for both DSI and DTI networks, and these differences were significant in both cases, although modestly more robust in DTI (two-tailed t-test, \( t = 6.15, P = 3e^{-6} \)) than DSI (\( t = 2.57, P = 0.017 \)) networks; see Fig. 4C.

**Basic connectivity properties**

Given the sensitivity of raw connectivity matrices, we next asked whether basic connectivity properties, which are global measures derived from those matrices, also displayed sensitivity to individual variability. As shown in Fig. 4D, DSI and DTI networks had vastly different weighted network properties; nodal strength, nodal diversity, and edge diversity were all higher in DSI networks, indicating not only that more tracts were reconstructed in DSI networks (as indicated by increased nodal strength) but also that the variability over brain regions (nodal diversity) and tracts (edge diversity) was higher in DSI networks.

Using the intra-class correlation coefficient (ICC), we specifically tested whether basic connectivity properties including the average weight, nodal diversity, edge diversity, and density showed more variability between subjects than within subjects over scanning sessions (see Materials and methods). While the reproducibility of density was fairly poor, as shown in Fig. 4E, the reproducibility of weighted connectivity measures was high (mean ICC=0.72) for all networks reconstructed from both DSI and DTI data.

We next sought to assess the relative effects of imaging modality (DSI or DTI), atlas and resolution on the reproducibility of weighted connectivity properties. A repeated measures ANOVA where modality (DSI or DTI), atlas and resolution were treated as a categorical factor and atlas and resolution were treated as repeated measures. Significant main effects were followed by posthoc testing, for which results are also reported.

### Table 2

<table>
<thead>
<tr>
<th>Measure</th>
<th>ICC</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Effect</strong></td>
<td><strong>Statistic</strong></td>
<td><strong>P-value</strong></td>
</tr>
<tr>
<td><strong>Modality</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posthoc: DSI vs DTI</td>
<td>( F=3.19, 6e^{-4} )</td>
<td>( F=345.77, 4e^{-5} )</td>
</tr>
<tr>
<td>Posthoc: AAL vs HO</td>
<td>( F=4.08, 0.113 )</td>
<td>( F=5.26, 0.083 )</td>
</tr>
<tr>
<td>Posthoc: HO vs LPBA40</td>
<td>( F=149.83, 2e^{-4} )</td>
<td>( F=143.35, 2e^{-4} )</td>
</tr>
<tr>
<td>Posthoc: AAL vs LPBA40</td>
<td>( F=218.15, 1e^{-4} )</td>
<td>( F=123.45, 3e^{-4} )</td>
</tr>
<tr>
<td><strong>Resolution</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posthoc: Original vs Sub 1</td>
<td>( F=14.31, 0.019 )</td>
<td>( F=7.90, 0.048 )</td>
</tr>
<tr>
<td>Posthoc: Sub 1 vs Sub 2</td>
<td>( F=12.53, 0.024 )</td>
<td>( F=18.65, 0.012 )</td>
</tr>
<tr>
<td>Posthoc: Sub 2 vs Sub 3</td>
<td>( F=7.27, 0.054 )</td>
<td>( F=10.34, 0.032 )</td>
</tr>
<tr>
<td><strong>Atlas × modality</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution = modality</td>
<td>( F=7.86, 0.003 )</td>
<td>( F=0.98, 0.431 )</td>
</tr>
<tr>
<td>Atlas × resolution</td>
<td>( F=1.75, 0.152 )</td>
<td>( F=2.15, 0.083 )</td>
</tr>
<tr>
<td>Atlas × resolution × modality</td>
<td>( F=2.11, 0.088 )</td>
<td>( F=1.27, 0.016 )</td>
</tr>
</tbody>
</table>

### Anatomical location of connectivity

The results up to this point have assessed global connectivity measures and their modulation by spatial resolution and imaging modality (DSI/DTI). We now build on these global findings by mapping the anatomical location of reconstructed tracts and assessing their reproducibility. Fig. 5 shows the nodal strength (or average number of tractography streamlines emanating from a brain region) of all regions in the DSI (left) and DTI (right) networks (averaged over subjects and sessions) for the 3 whole-brain atlases (AAL, top; HO, middle; LPBA40, bottom).

Several important results are evident from these visualizations. Perhaps the most striking is that connectivity appears to be strongly dependent on the atlas used, and regional connectivity appears higher for the larger regions in each atlas. Indeed, the size of a region (in \( \text{mm}^3 \)) was highly correlated with the strength of a region: Pearson's correlation coefficient \( R = 0.78 \) and \( P < 4e^{-16} \) (see Supplementary Fig. 12 for scatterplots over all atlases and resolutions). An immediate correlate of this result is that hub distributions in upsampled templates will be qualitatively similar to those in their mother atlas (Supplementary Figs. 10–11). Importantly, the dependence of connectivity on region size indicates a degree of homogeneity in cortical connectivity structure, where the measured connectivity is highly predicted by the size of the sample whose connectivity is being measured.

Beneath each mapping of nodal strength provided in Fig. 5.5 is a mapping of the reproducibility (ICC) of that nodal strength over the 3 sessions. Across atlases and modalities, the strength of middle frontal and central cortices is largely reproducible while the strength of inferior temporal and occipital cortices is less reproducible.

### Variability and reproducibility of network architecture

In a final branch of analysis, we determined the variability and reproducibility of measured network architecture using binary network thresholds over a range of network densities.

### Reproducibility of graph metrics

Graph metrics are inherently diverse in definition, variability, and sensitivity. It is therefore intuitive that variability and reproducibility might vary from metric to metric, as is indeed evident from Fig. 6. The most reproducible metrics retained information regarding the physical length of connections. In DSI, the mean connection distance was the most reproducible (ICC = 0.64), while in DTI networks the Rent's exponent was the most reproducible (ICC = 0.70); see
Supplementary Table 3 for a list of all graph metrics in order of reproducibility. Interestingly, the thirteen binary graph metrics used in this study were less reproducible on average than basic connectivity properties as measured by a two-tailed t-test of their ICC values (t = 7.40, P = 2e−9); see Fig. 6A.

In order to distill the relative effects of varying methodological parameters, we performed a repeated measures ANOVA on the ICCs of the graph metrics, where modality (DSI vs DTI) was treated as a categorical factor, and atlas and resolution were treated as repeated measures; see Table 3. This ANOVA showed a significant main effect of modality where posthoc comparisons confirmed that graph metrics derived from DSI networks were less reproducible than those derived from DTI networks. The atlas applied to the brain had a significant effect on reproducibility and posthoc comparisons showed that the AAL atlas produced more reproducible results than either the HO or the LPBA40. The ANOVA also showed a significant main effect of resolution, and posthoc comparisons showed that the LPBA40. The ANOVA also showed a significant main effect of resolution, and posthoc comparisons showed a steady decrease in reproducibility with increasing resolution (the original atlas is significantly more reproducible than the densest upsampling, Sub 3 but no other pairwise comparisons were significant). Interestingly, this result highlights differential effects of spatial resolution on basic connectivity and graph properties: basic connectivity properties increase in reproducibility with increasing spatial resolution while graph metrics decrease in reproducibility.

A similar repeated measures ANOVA on the coefficient of variation (Fig. 6B) revealed no significant effects, suggesting that the variability of graph metrics was not significantly affected by modality, atlas, or spatial resolution; see Table 3. However, it is important to note that the CV of several graph metrics is consistently lower in DSI networks compared to DTI, indicating a potential advantage of DSI networks in producing reliable estimates of graph parameters with very little variability. In general, CV values were close to 1%, being smallest for the metric of robustness to random attack and highest for the assortativity. However, it is important to note that the CV as a statistical measure is sensitive to low values of the mean, which is particularly true for the assortativity measure especially in DSI networks, which may explain the unusually high CV value for this metric.

**Alternative weighting schemes**

The definition of edge weight that we have used in this work is the number of tracts that have been reconstructed traveling from any node i to any node j. It is important to be cautious in interpreting tract counts as a direct measurement of the absolute volume of connectivity between two regions, largely because several important factors may artifically inflate or deflate tract counts. These factors include crossing fibers, fiber length, and proximity to large tracts like the corpus callosum, each of which may not be homogenous throughout the brain. However, it is important to note that this definition of edge weight is not the only one that is available to the investigator. Tract counts may be weighted by the size of a region to create a tract density, or by the length of tracts as an attempt to correct for any distance bias in tractography algorithms (Hagmann et al., 2008). Even more broadly, edge weights may be based on alternative variables such as average FA or tract curvature.

While it was not the purpose of the present study to provide an exhaustive account of the effect of weighting scheme on reproducibility, we did consider the effects of a single alternative weighting scheme in which the size of a cortical region was no longer a predictor of its connectivity. In Fig. 7A, we show that after regressing out the effect of region size for all three classical atlases, high strength nodes or ‘hubs’ are located in lateral regions adjacent to the central sulcus.
graphs such as those we analyzed above in the first section of the Results, where we identified conserved architectural principles such as high clustering, short pathlength, assortativity, hierarchy, and small-worldness. However, weighting schemes may have an impact on network measures derived from thresholded graphs at any particular network density or averaged over a range of network densities. To assess the extent of this impact, we redefined the edge weight between any two regions by dividing the total tract count by the sum of the sizes of the two regions that it connects (Hagmann et al., 2008). In Fig. 7B, we see that the value of some graph measures, averaged over network densities, may be modulated by this alternative weight definition (pathlength), while others may not (clustering). However, the variability and reproducibility of graph metrics are on average unaffected by the choice of weighting scheme; see Fig. 7C.

**Discussion**

In this study, we demonstrated the presence of robust architectural principles in measured networks of white matter connectivity which are independent of modality (DSI and DTI), atlas (AAL, HO, and LPBA40), and spatial resolution (for networks ranging in size from 54 to 880 nodes). These highly conserved properties include sparsity, small-worldness, hierarchy, assortativity, and several measures of topo-physical interdependence including the property of Rentian scaling. We further assessed the ability of both basic connectivity properties and graph theoretical properties to successfully measure individual variability in brain structure. We found that weighted connectivity properties of networks derived from both DSI and DTI data were consistently reproducible (ICC=0.72) and displayed low variability (CV~5%), suggesting their potential usefulness in clinical applications. While reproducibility of network properties was higher in DTI, DSI networks were characterized by an increased number of reconstructed tracts, as well as a higher diversity of that number over available cost regime. Properties in (A) and (B) are listed along the x-axis in the following order: global connectivity (pathlength, global efficiency, and betweenness), local connectivity (clustering and local efficiency), global-local properties (modularity and hierarchy), dynamics (synchronizability), mixing (assortativity), robustness (robustness to targeted attack and robustness to random attack), topo-physical (Rent's exponent), and physical (mean connection distance). Error bars indicate standard deviation in the mean over atlases and resolutions.

and perisylvian fissure as well as in medial regions including the precuneus, cingulate, and medial frontal cortex. Note that after regression of region size, the hub distributions are significantly more similar across atlases than was seen previously (Fig. 5).

Following this analysis, it is natural to ask whether alternative weighting schemes have an impact on measured network architecture or its reproducibility. We note that weighting schemes do not have any impact on binary graph metrics derived from unthresholded

### Table 3

<table>
<thead>
<tr>
<th>Measure</th>
<th>ICC</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posthoc: DSI vs DTI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atlas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posthoc: AAL vs HO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posthoc: AAL vs LPBA40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posthoc: HO vs LPBA40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posthoc: original vs Sub 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution × modality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atlas × resolution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atlas × resolution × modality</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 6.** Individual variability and reproducibility of graph architecture. The intra-class correlation coefficient (ICC) (A) and coefficient of variation (CV) (B) of graph metrics for both DSI (red) and DTI (black) networks, averaged over subjects and scans. The ICC and CV are each computed for the value of each network metric averaged over the available cost regime. Properties in (A) and (B) are listed along the x-axis in the following order: global connectivity (pathlength), global efficiency, and betweenness), local connectivity (clustering and local efficiency), global-local properties (modularity and hierarchy), dynamics (synchronizability), mixing (assortativity), robustness (robustness to targeted attack and robustness to random attack), topo-physical (Rent's exponent), and physical (mean connection distance). Error bars indicate standard deviation in the mean over atlases and resolutions.
and DTI), and parcellation atlas. All reconstructed networks were characterized by a sparse connectivity between nodes, a property which has consistently been reported for anatomical networks derived from both direct and indirect measurements in humans (Bassett et al., 2008; He et al., 2007; Hagmann et al., 2008; Gong et al., 2009), macaque (Fellows and Van Essen, 1991; Young et al., 1995; Hilgetag et al., 2000), cat (Scannell et al., 1995, 1999), and C. elegans (White et al., 1986; Kaisera and Hilgetag, 2006; Choe et al., 2004). Such sparse connectivity is thought to be caused by an evolutionary pressure for energy efficiency (Attwell and Laughlin, 2001; Niven and Laughlin, 2008), where connections cluster in brain regions where integrated information processing must occur, and only sparsely link these clusters, creating a modular structure (Sporns et al., 2000; 2004; Bassett and Bullmore, 2006). The robust presence of network sparsity over multiple spatial resolutions (e.g., with N varying from 54 to 880) is a necessary prerequisite for non-random topological structure over a broad range of spatial scales; see Fig. 2A.

Conserved topological relationships

As has been consistently observed in functional and anatomical brain networks in a wide variety of species, all networks studied in this work displayed distinctly non-random topological properties including a relatively high clustering coefficient and a relatively short pathlength (Bassett and Bullmore, 2006; Bullmore and Sporns, 2009); see Fig. 2B. The combination of these two non-random properties, known as small-worldness, is robust to atlas choice in functional brain networks (Wang et al., 2009) and robust to spatial resolution in both anatomical and functional networks (Zalesky et al., 2010; Hayasaka and Laurienti, 2010). The term ‘small-world’ was coined several decades ago in the context of the social phenomenon of friendship (Milgram, 1967; de Sola Pool and Kochen, 1978), and the property has since been suggested to confer a degree of efficiency in information transfer through a system (Barahona and Pecora, 2002). However, the presence of the small-world property in systems with very disparate network architectures suggests that it alone cannot fully describe the details of graph organization.

Therefore, it is important to note that assortative degree mixing, or ‘assortativity’, was also consistently displayed in the human cortical networks studied here; see Fig. 2D. Assortativity, a measure of the preference for a node to connect to other nodes with the same degree, has been previously reported for a subnetwork of the human brain (Bassett et al., 2008), and for whole-brain networks derived from DSI data (Hagmann et al., 2008). Our current results both confirm and extend these findings to whole-brain networks derived from DTI data. In comparing this result to other biological systems, it is interesting to note that the canonical neuronal network of C. elegans has been previously demonstrated to be disassortative rather than assortative (Newman, 2002). In fact, a variety of biological networks including food webs and protein interaction networks are broadly disassortative while social networks such as coauthorship and collaboration networks have been found to be assortative (Newman, 2002). The unique character of these white matter connectivity profiles as assortative biological networks suggests that this metric in particular has some discriminatory power, and this result may give us further insight into appropriate neurodevelopmental paradigms for modeling growth and development of these systems.

A complementary measure, the property of topological hierarchy, where network hubs tended to be less clustered than non-hubs indicating their ‘connector’ rather than ‘provincial’ role in the network (Ravasz and Barabasi, 2003; Guimera et al., 2007), was also consistently expressed in these networks; see Fig. 2C. Topological hierarchy directly impacts network dynamics (Zhou et al., 2006; Muller-Linow et al., 2008) by allowing for a structured segregation of function into small modules while that function is in turn governed by nodes at higher levels of the hierarchy. Hierarchy is common in corporate, information, social and language networks (Ravasz and Barabasi, 2003), and has also been reported in a subnetwork of heteromodular association cortices in the healthy human brain (Bassett et al., 2008). Interestingly, in the original derivation and application of the hierarchical measure (Ravasz and Barabasi, 2003), networks which were not physically embedded showed hierarchical structure.
while networks which were embedded lacked hierarchical structure. However, as we see here, some physically embedded systems can display robust hierarchical structure perhaps even due to their physical constraints on wiring (Bassett et al., 2008).

This combination of results, i.e., that the anatomical networks are characterized by sparsity, high clustering, short pathlength, assortativity, and hierarchy, could form the topological basis for the hierarchically modular structure of anatomical connectivity that has been both theoretically expected (Simon, 1962) and empirically described (Felleman and Van Essen, 1991; Hubel and Wiesel, 1961) for several decades. Recently, hierarchically modular network topology has been shown to exist in a wide range of biological networks (Ravasz, 2009; Reid et al., 2009) as well as information processing systems like the human brain and complex computer circuits that have been cost-efficiently embedded into physical space (Bassett et al., 2010).

Physical embedding of topology

We in fact find that there is a strong coupling between topological structure and physical structure in the human brain, consistent with previous results in other species (Costa et al., 2007). Perhaps the simplest display of the complementary structure in topological and geometric space is the striking relationship between topological distance (pathlength) and physical distance (Euclidean distance between center of mass of regions); that is, two regions which are far from each other in physical space are close to each other topologically (see Fig. 3). While the topological distance is a measure of global connectivity, it is interesting to note that this interdependence between space and topology extends to local connectivity, where regions clustered in topological space are more likely to be clustered in physical space, suggesting an inherently physical basis for anatomical (Bassett et al., 2010) and by extension functional (Meunier et al., 2009a,b) modularity. The combination of these two properties is compatible with the known importance of geographical constraints imposed on the brain both through evolution and during neurodevelopment (Durbin and Mitchison, 1990; Chklovskii et al., 2002; Chklovskii, 2004).

Geographical considerations can fundamentally constrain a network, e.g. through the physical size of the cortex determined at birth, or they may restrict the development of that network, for example, by imposing a metabolic cost on the development of long, myelinated axons (Durbin and Mitchison, 1990; Chklovskii et al., 2002; Chklovskii, 2004; Niven and Laughlin, 2008). Such a selection pressure for efficiency of cortical wiring may explain the robust presence of Rentian scaling in these networks, a property originally discovered in very large-scale integrated computer circuits (Christie and Stroobandt, 2000; Ozaktas, 1992), and characterizing efficient physical embedding of the system. While Rentian scaling has been reported to exist in a small sample of DSI networks for a single template and resolution (Bassett et al., 2010), here we report for the first time the robust presence of this phenomena across imaging modality (DSI or DTI), atlas, and spatial resolution; see Fig. 3D.

Modulation of measured architecture by size and anatomical parcellation

While the density of connections was consistently sparse in these anatomical networks, it was also modulated by spatial scale, dropping off exponentially with the number of nodes in the parcellation applied to the brain, and therefore the number of nodes in the network (see Fig. 2A and Supplementary Materials). While it has been previously suggested that density and the number of nodes are related (Zalesky et al., 2010), we have further extended this finding to multiple parcellation schemes and imaging modalities in addition to mathematically characterizing the form of that relationship as a power-law with consistent exponents in both DSI and DTI networks. Power-law relationships between measures of total connectivity (edges) and measures of total processing blocks (nodes) are a well-known phenomenon in the very large-scale integrated circuit literature (Bakoglu, 1990) and indicate a consistent organizational structure over variations in system size. In the current context, this finding that cortical white matter connectivity has architectural properties which are conserved over a range of scales, is consistent with similar recent findings in human brain functional networks (Bassett et al., 2006, 2009; Kitzbichler et al., 2009).

In addition to density, the majority of network metrics studied here also varied with spatial resolution to different degrees: some strongly (e.g., exponentially, due to a strong dependence on density) and some weakly (see Fig. 2). For example, as reported for DTI networks previously (Zalesky et al., 2010), the clustering coefficient and pathlength both increased with increasing spatial resolution. It may be that this modulation of structural topology by spatial resolution underlies apparent alterations in functional topology at varying spatial scales, due to the complex structure–function relationship (Honey et al., 2007, 2009, 2010). Our results may therefore partially explain the reported dependence of degree distributions in functional networks on spatial resolution (van den Heuvel et al., 2008; Egilmez et al., 2005; Hayasaka and Lauriello, 2010), where larger hubs are more likely to occur at higher spatial resolutions.

Because density varied with the number of nodes in the network, we cannot dissociate the effects of these two measures of network size on graph metrics values. It is interesting to note, however, that the amount of variation was metric-dependent; modularity, hierarchy, assortativity, synchronizability, and mean connection distance were more stable across spatial resolutions than the other graph metrics studied here. It has been previously reported that graph metrics are independently sensitive to variations in density (Anderson et al., 1999), and until we better understand these inter-dependencies, holding the number of nodes and edges fixed in an experimental analysis will likely remain a common practice. It is as yet not well understood what sorts of organizational structures could display this combination of sensitivity of some graph metrics to spatial scale and insensitivity of others, and this will likely be an important and interesting area of future research.

Individual variability of measured architecture

Our results showed that connectivity matrices derived across both individuals and scans were highly similar in both imaging modalities, suggesting a robust conservation of human connectivity diagrams. Our work did, however, uncover strong differences between the coverage and sensitivity of DSI and DTI networks. DSI networks were characterized by $\sim 1.1 \times$ greater density, $\sim 3.5 \times$ greater average weight, $\sim 3.0 \times$ greater nodal diversity, and $\sim 3.6 \times$ greater edge diversity than their DTI counterparts. This modality dependent extent of connectivity can be understood in the context of the inherent differences between the two imaging techniques. DSI data is acquired using an order of magnitude more diffusion directions and can successfully resolve the crossing of fiber tracts (Wedeen et al., 2005), providing a significantly more complete wiring diagram (Hagmann et al., 2008). Possible scientific questions in which DSI acquisitions may therefore be preferred include those which would benefit from more extensive cortical coverage and those that would require sensitivity to the diversity of connectivity profiles over cortical regions (for example, in schizophrenia (Lynam et al., 2010)).

Furthermore, intra-subject similarity was in general higher than inter-subject similarity, arguably indicating that these analysis methods highlight biologically relevant individual variation in connectivity structure. The fact that intra-subject and inter-subject similarities are less different in DSI than in DTI matrices may be due to their relatively higher connectivity. It is intuitively plausible that a few
individual signatures of connectivity in the dense DSI matrices have a smaller statistical impact on the matrix similarity than the same number of signatures in the sparser DTI matrices.

The subsequent results for reproducibility of both basic connectivity properties and graph metrics highlight a potential trade-off between accuracy and reproducibility in diffusion imaging networks: while DSI networks are arguably more comprehensive, containing more tracks, and more successfully resolving the crossing of tracts, they are also less reproducible. On the other hand, DTI networks are highly reproducible but provide less information, being sparse and containing fewer tracts. Despite these dramatic differences in network constituents, it is impressive to note that both types of networks consistently displayed a large range of conserved architectural characteristics, as discussed above. Further, it is important to note that while higher for DTI networks, reproducibility of weighted connectivity properties was still significant in all DSI networks studied (ICC = 0.72), indicating that measures derived from either modality can be successfully applied to studies of individual variation in cognitive ability, clinical or drug studies, or studies requiring multiple scanning sessions.

A second important trade-off between variability and reproducibility is evident from our results. The coefficient of variation, which measures the variability of metrics, was not significantly different between graph metrics derived from DTI and DSI networks, while the intra-class correlation coefficient, which measures the differences between individuals compared to the differences between scans, was significantly different. In combination, these results suggest that the DSI networks provide a robust estimate of structure that is consistent across individuals but are less sensitive to differences over time within an individual. Examining these results from a different perspective, we note that it is potentially possible that global graph metrics, due to their agglomerative nature, are relatively less sensitive to small changes in network structure in more homogeneously connected graphs (DSI networks) than in less connected graphs (DTI networks). If this were true, it would suggest that assessing regional network statistics or using weighted network measures may enable more robust estimation of individual differences in DSI than DTI networks. These and other possibilities need to be addressed systematically, and we are actively pursuing these research questions.

In addition to differences in connectivity and properties of global graph metrics, there are several other potential reasons for the decreased reproducibility evident in the DSI networks in this study primarily related to acquisition and preprocessing. Firstly, it is important to note that the DSI acquisition is 5 times longer than the DTI acquisition (45 min as opposed to 9 min), increasing the probability of significantly more cumulative head movement at the conclusion of the DSI scan. Furthermore, movement correction in the current acquisition sequence is difficult because the b-values of each scan increase with scan duration, and high b-value scans not only have a lower signal-to-noise ratio but also have a larger number of diffusion directions. More extensive movement correction may be possible in alternative scanning sequences where multiple b0 scans are acquired throughout the protocol or by decreasing total scan duration using the technique of simultaneous image refocusing echo-planar imaging (Reese et al., 2009).

Preprocessing steps including the brain parcellation and diffusion tractography may also have a nontrivial impact on the reproducibility of connectivity structure. The choice of tractography algorithm could theoretically affect the reproducibility of DSI network structure for example by altering the estimation of the ODF peaks, the maximum number of ODF peaks detected per voxel, and the propagation through fiber crossings. Furthermore, in deterministic tractography, noise artifacts increase with streamline length, a relationship which may differentially impact reproducibility in DSI networks which have longer tracts than DTI networks. Secondly, the application of whole-brain atlases like the HO, AAL, and LPBA40 to individual subjects and scans necessarily blurs the boundaries between individuals. Because DSI networks are composed of more tracts, and more spatially distributed tracts, they may be more sensitive to poor inter-subject concordance of brain regions. In order to retain more individual information regarding cortical structure, a subject-specific cortical surface reconstruction and parcellation, such as that provided by FreeSurfer (Destrieux et al., 2010), could be performed.

In light of this discussion, it is important to note that it was not the purpose of this particular study to determine the optimal processing pathway to create the most robust human brain white matter connectivity atlas. Instead the relative power of our approach lies in its simplicity. Using freely available software and minimal adaptation, we have uncovered a range of conserved network architectural principles which are robust across multiple scanning sessions to variability in imaging technique, whole-brain atlas, and spatial resolution, and which we therefore have reason to expect will also be robust to alternative processing streams. It is, however, possible that fine tuning the current approach or using alternative approaches may significantly increase reproducibility of connectivity structure in human brain anatomical networks derived from both DTI and DSI data.

Graph metrics

Our results show that network metrics were differently sensitive to individual variation, as measured by the ICC of network metrics over scans; see Fig. 6. Graph properties derived from both DSI and DTI networks varied largely in reproducibility, with ICC values ranging from 0.30 to 0.70, similar to those reported for functional MEG networks (Deuker et al., 2009). The small variation in reproducibility with methodological variation is consistent with a recent study showing that reproducibility of topological measures in DTI networks is largely independent of gradient scheme, number of diffusion directions, and gradient amplitude Vaessen et al. (2010). Interestingly, the most reproducible network metrics, across atlases, resolutions, and modalities (DSI and DTI), were those which retained information regarding the physical extent of connections. In DSI networks, the mean connection distance was the most reproducible while for DTI networks, both the mean connection distance and the Rent’s exponent showed the highest ICC values. Such metrics may retain proportionally more power than purely topological metrics by combining individual variation in both space and topology (Bassett et al., 2010). The fact that basic connectivity properties were more reproducible than graph properties may be due to their relative mathematical simplicity, based on simple means and variances, which could make them robust to alterations in the number of edges. Conversely, graph properties measure more complicated relationships between edges and the nodes they connect, and therefore may be both more sensitive and more fragile to edge alterations.

We finally provide preliminary evidence for the modulation of graph metric values by and the robustness of variability and reproducibility metrics to alternative weighting schemes. However, further work is necessary to systematically assess the effects of multiple weighting schemes, including those with length bias corrections (Hagmann et al., 2008), on network architecture and its reproducibility, work which is unfortunately outside of the scope of the present study.

Anatomy

In addition to global connectivity, networks are often assessed for anatomical distribution of connectivity from poorly connected ‘provincial’ nodes to highly connected ‘hubs’ (Albert and Barabási, 2002; Amaral et al., 2000). These hubs likely support the majority of network traffic, and so their identity and characterization is important in describing the robustness and sensitivity of the network architecture (Albert and Barabási, 2002; Albert, 2005). In brain graphs, functional hubs have been found predominantly in association cortex
In anatomical networks, hubs have also been reported in association areas (He et al., 2007), but Zalesky et al. (2010) suggest that this result may depend on the spatial resolution of the atlas used to parcellate the brain. Our results uncover a strong relationship between region size and strength, such that “hubs” are predictably the largest brain regions in the atlas; see Fig. 5. The strength of this relationship indicates a degree of homogeneity in cortical connectivity such that the number of tracts measured in a given region depends strongly on the size of the measured region. A similar result has in fact been reported for tract-tracing studies in macaque monkeys, where the strength of a connection originating from a cortical area was found to be positively correlated with the size of the region’s surface (Hilgetag and Grant, 2000).

However, an important corollary of this finding for the current study is that highly connected nodes identified using raw connectivity weights cannot be interpreted as ‘hubs’ in the common network understanding of the word. Instead, hubs are those regions which are more highly connected than might be expected due to their size, and these occur predominantly in regions adjacent to the central sulcus and perisylvian fissure as well as regions along the medial wall largely associated with the functional default mode network of the resting state brain (Grecius et al., 2003), consistent with previous work (Hagmann et al., 2008).

In addition to the anatomical localization of connectivity, we mapped the reproducibility of that connectivity across the surface of the cortex. In both DTI and DSI networks, we found that regions with poorly reproducible connectivity strength were consistently located in the inferior temporal and occipital cortex while regions with highly reproducible connectivity strength were located in the middle frontal and distributed medial cortices; see Fig. 5. The anatomical localization of reproducibility suggests that it is driven at least in part by magnetic susceptibility (Vargas et al., 2009).

**Methodological limitations**

**Atlas and upsampling**

There are several methodological limitations to the current work, perhaps the simplest of which being that we have not performed an exhaustive analysis of all possible whole-brain atlases. We have instead chosen to focus on three commonly used and representative atlases (AAL, HO, and LPBA40) which span a range of spatial scales (from the 54 regions of the LPBA40 atlas to the 110 regions of the HO atlas) and which further provide alternative parcellations to multiple brain regions including large portions of the frontal, temporal, and occipital cortices. We have also used a uniform upsampling of these atlases in order to retain the original cytoarchitectonic boundaries of the mother atlas, and therefore our results do not address the effects of random upsampling, non-uniform upsampling, or upsampling performed without respecting anatomical boundaries (Zalesky et al., 2010).

**Tractography and acquisition**

Construction of a human connectome depends upon tracking individual neuronal fiber tracts through diffusion-based images. Such tracking can be performed using a wide variety of methods including streamline tracking such as deterministic tractography (Basser et al., 2000; Conturo et al., 1999; Mori and Barker, 1999; Lazar et al., 2003) probabilistic tractography (Behrens et al., 2003; Hagmann et al., 2003; Parker and Alexander, 2003), front evolution tracking (Parker et al., 2002; Campbell et al., 2005), and graph-based tracking (Iturria-Medina et al., 2007, 2008; Sotiropoulos et al., 2010). We restricted our attention in the current work to deterministic streamline tracking methods. Deterministic algorithms generate tracts using locally optimal decisions while probabilistic algorithms generate continuous probabilities of a tract’s existence. Both methods have had many successes, but also come with some limitations (Behrens et al., 2007; Sherbondy et al., 2008). Deterministic tractography is both simple and swift, and produces discrete and therefore easily interpretable results and is simply applicable to both DTI and DSI acquisition streams. However, probabilistic algorithms have the advantage of assessing a wider pathway search space, and therefore an important future study will focus on reproducibility of DTI network structure as a function of tractography algorithm (Moldrich et al., 2010). Determining the best algorithms for both deterministic and probabilistic tractography is still a growing field of important research (Behrens et al., 2007; Descoteaux et al., 2009; Sherbondy et al., 2008), and current efforts are complicated by false positives, false negatives, and biases against long-distance connections. Future advances in tractography algorithms will therefore likely provide a better springboard for further reproducibility studies of this kind. While a systematic study of preprocessing strategies is outside the scope of this paper, it would theoretically be interesting to address in future whether reproducibility can be increased by altering the acquired slice thickness, using more or less diffusion directions, using a higher channel coil, and alternating between low and high b-values during the acquisition.

**Network metrics**

The current network analysis described in this paper has focused on quantifying the reproducibility of global network properties using a combination of basic connectivity and binary graph metrics, the most commonly applied methods to date (Gong et al., 2009; Vaessen et al., 2010; Shu et al., 2009; Li et al., 2009; Zalesky et al., 2010). However, additional work could also include an assessment of weighted network measures in addition to more detailed quantification of regional reproducibility (Hagmann et al., 2008; Iturria-Medina et al., 2007). While it was not the subject of the current work, it would be interesting in future to determine whether or not alternative thresholding algorithms based on local thresholds or minimum spanning trees (Hagmann et al., 2008) might provide better reproducibility.

**Conclusion**

In this work, we have explored the topological basis of hierarchical modularity using metrics which are consistently expressed in measured cortical architecture across multiple alternative mappings. Combined with conserved topo-physical relationships, these results point to a highly structured and physically-dependent nervous system. Our work also highlights specific network properties, such as Rent’s exponent, that are particularly sensitive to individual variation in anatomical connectivity while remaining insensitive to inter-scan variability. We demonstrate that the network connectivity architecture of both DSI and DTI networks is highly reproducible. Because both of these imaging acquisition techniques and subsequent tractography algorithms are still being developed, our understanding of canonical cortical architecture is likely to deepen in future. Based on our work, however, we suggest that network analysis of human white matter connectivity can currently provide sensitive topological and physical measurements of cortical structure across multiple temporal and spatial scales.

**Acknowledgments**

This work was supported by the David and Lucile Packard Foundation, PHS Grant NS44393 and the Institute for Collaborative Biotechnologies through contract no. W911NF-09-D-0001 from the U. S. Army Research Office. J.A.B. was supported by an NIH NRSA, grant number 1F31AG035438-01.