SCALAR CONNECTIVITY MEASURES FROM FAST-MARCHING TRACTOGRAPHY REVEAL HERITABILITY OF WHITE MATTER ARCHITECTURE

Vishal Patel,[1] *Ming-Chang Chiang,*[1] *Paul M. Thompson,*[1] *Katie L. McMahon,*[2] *Greig I. de Zubicaray,*[2] *Nicholas G. Martin,*[3] *Margaret J. Wright,*[3] *Arthur W. Toga*[1]

[1] Laboratory of Neuro Imaging, University of California, Los Angeles
[2] Functional Magnetic Resonance Imaging Laboratory, University of Queensland, Brisbane
[3] Queensland Institute of Medical Research, Brisbane

ABSTRACT

Recent advances in diffusion-weighted MRI (DWI) have enabled studies of complex white matter tissue architecture *in vivo*. To date, the underlying influence of genetic and environmental factors in determining central nervous system connectivity has not been widely studied. In this work, we introduce new scalar connectivity measures based on a computationally-efficient fast-marching algorithm for quantitative tractography. We then calculate connectivity maps for a DTI dataset from 92 healthy adult twins and decompose the genetic and environmental contributions to the variance in these metrics using structural equation models. By combining these techniques, we generate the first maps to directly examine genetic and environmental contributions to brain connectivity in humans. Our approach is capable of extracting statistically significant measures of genetic and environmental contributions to neural connectivity.

Index Terms— Magnetic resonance imaging, genetics, nervous system, algorithms, brain

1. INTRODUCTION

The differential contributions of genetic and environmental factors to the patterns of structural connectivity in the central nervous system are topics of great interest and importance. Structural (T1-weighted) MRI studies have demonstrated that a number of morphometric features of the brain, such as cortical thickness and regional gray and white matter volumes, are heavily influenced by genetics. Few studies, however, have examined genetic effects on the connectional architecture as probed through diffusion-weighted MRI (DWI). Recent advances in this modality have enabled the reconstruction of a variety of intravoxel probability distributions and metrics to represent local white matter microstructure. The diffusion tensor (DT) model estimates a single Gaussian diffusion profile at each voxel [1]. Other advanced methods attempt to resolve the orientational or radial structure of spin diffusion, or the distribution of the underlying fibers themselves [2, 3, 4].

This work was supported by the National Institutes of Health through grants 1U54RR021813-01 (Center for Computational Biology), 5T32GM008042-25 (Medical Scientist Training Program), and R01 HD050735, and by the NHMRC, Australia.

Quantifying connectivity, however, requires that the information on these structures be extended across voxel boundaries to reconstruct anatomical fiber pathways. The few existing diffusion imaging studies addressing genetic and environmental effects have focused on intravoxel measures of fiber integrity (fractional anisotropy) [5, 6] or local measures of fiber complexity (Jensen-Shannon divergence) [6]. In this report, we seek to quantify long-distance structural connectivity rather than local architectural features and derive a new set of measures based on a fast-marching tractography algorithm. Several groups have recently suggested fast-marching as a suitable method to extract connectional pathways from DWI [7, 8, 9]. Fast-marching presents a much lighter computational burden than the probabilistic class of tractography algorithms [10], an important consideration for genetic studies that require the evaluation of large numbers of subjects in order to attain statistical power. Moreover, unlike conventional streamline methods [11], the fast-marching approach reports information about the validity of each individual trajectory—an essential concern for any quantitative analysis of connectional heritability.

Below, we outline the details of the fast-marching algorithm and associated connectivity metrics. We generate connectivity maps for a real DTI dataset from 92 twins and analyze the variance of these metrics using structural equation modeling, a statistical method that decomposes the variance in observed quantities into contributions from a set of latent variables—in this case, genetic and environmental factors. By combining these techniques, we generate the first maps to directly examine genetic and environmental contributions to brain connectivity in humans, and we further demonstrate that these maps provide statistically significant measures of connectional heritability.

2. METHODS

2.1. Data acquisition and preprocessing

23 pairs of monozygotic twins (MZ; 11 male pairs/12 female pairs; age = 25.1$\pm$1.5 years) and 23 pairs of dizygotic twins (DZ; all

same-sex pairs; 10 male pairs/13 female pairs; age = 23.5±2.1 years) were recruited from different families and scanned at the Center for Magnetic Resonance at the University of Queensland using a 4 Tesla Bruker Medspec scanner with a transverse electromagnetic (TEM) headcoil. Diffusion-weighted scans were acquired using single-shot echo-planar imaging with a twice-refocused spin echo sequence to reduce eddy-current induced distortions. Acquisition parameters were optimized to provide the best signal-to-noise ratio for estimation of DTs. Imaging parameters were: 21 axial slices (5 mm thick), FOV = 23 cm, TR/TE 6090/91.7 ms, 0.5 mm gap, with a 128×128 acquisition matrix. 30 images were acquired: 3 with no diffusion sensitization ($\mathrm{b} = 0$, i.e., T2-weighted images) and 27 diffusion-weighted images ($\mathrm{b} = 1132\,\mathrm{s/mm^2}$) in which the gradient directions were evenly distributed on the sphere. The reconstruction matrix was 128×128, yielding a $1.8 \times 1.8\,\mathrm{mm}^2$ in-plane resolution. The total scan time was 3.05 minutes.

For each subject, DT images (denoted $D_{ij}, 1 \leq i, j \leq 3$) were computed using MedINRIA [`http://www-sop.inria.fr/asclepios/software/MedINRIA/`]. The D_{11} image was manually skullstripped, yielding a binary brain mask. The masked image was then registered to the ICBM53 average brain template with affine transformation using FLIRT [`http://www.fmrib.ox.ac.uk/fsl/flirt/`], and resampled to isotropic voxel resolution ($128 \times 128 \times 93$ voxels, 1.7 mm in size). The resulting transformation parameters were used to rotationally reorient the tensor at each voxel, and then affine-align the tensor-valued images based on trilinear interpolation of the log-transformed tensors. Tensors were rotated and translated but not scaled. All affine-registered DT images were then registered to a randomly-selected subject's image using an inverse-consistent fluid registration algorithm that minimizes the symmetrized Kullback-Leibler (sKL) divergence of the two tensor-valued images [12].

2.2. Fast-marching tractography

The fast-marching (FM) method is a means for solving a moving boundary problem by converting it to a stationary form initial value problem. We extended this idea to tractography by formulating an ordered upwind method, in an Eulerian framework, that solves the static convex Hamilton-Jacobi partial differential equation:

$$\begin{aligned} H\left(\nabla u\left(\mathbf{x}\right), \mathbf{x}\right) &= 1, \quad \mathbf{x} \in \Omega \subset \mathbb{R}^3 \\ u\left(\mathbf{x}\right) &= 0, \quad \mathbf{x} \in \partial\Omega \end{aligned} \tag{1}$$

where $H\left(\cdot,\cdot\right)$ is the Hamiltonian, $u\left(\mathbf{x}\right)$ represents the time at which the boundary passes point $\mathbf{x}$, and $\partial\Omega$ defines an initial boundary condition. Mathematical formalisms for ordered upwind methods are presented in [13]; here we focus on how to practically implement these ideas for tractography. The following description reviews the "advanced fast marching" approach taken by [8], with important differences for extracting connectivity metrics as noted.

For tractography, Eq. (1) can be understood as the problem of determining the time at which a front emanating from some seed region of interest (ROI) arrives at each voxel in the rest of the volume. The FM method dynamically assigns each voxel in the volume to one of three mutually exclusive groups: voxels that have already been *accepted* into the front, those in the narrow *band* neighboring the front, and those *outside* this band (Fig. 2.2). Initialization involves setting the seed voxels to have an arrival time of zero and labeling them *accepted*. The *band* is the set of all voxels in the 98-neighborhoods of voxels in the *accepted* group which are not themselves *accepted*. The 98-neighborhood contains all voxels in the adjacent $3 \times 3 \times 3$ set, plus voxels from the $5 \times 5 \times 5$ group whose displacement vectors are not collinear with those in the smaller set. Each voxel in the *band* is assigned an arrival time according to a speed function described further below. At each iteration, the *band* voxel with the earliest arrival time is moved to the *accepted* group and the *band* is updated accordingly. Termination occurs when all voxels have been *accepted*.

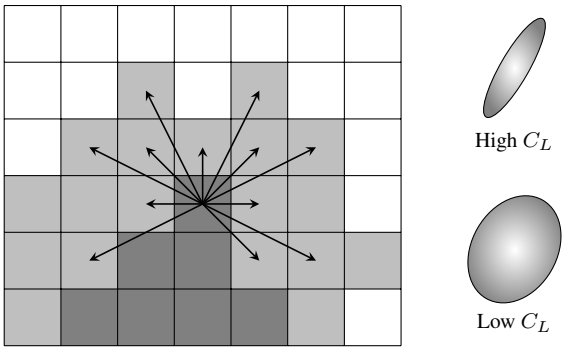

Fig. 1. Front evolution in fast-marching tractography. One voxel from the band of consideration (light gray) is accepted into the advancing front (dark gray) at each iteration. Arrows indicate the potential transitions from one voxel on the edge of the front. Voxels beyond the band (white) are not considered. Selection of which transition to make is based on a speed function which partially depends on the C_L of the diffusion tensors at the source and destination voxels, illustrated at right (see text for definition of C_L).

The FM front evolution is guided entirely by the choice of speed function. Here, we select one of four speed functions based on the linear anisotropy ($C_L = (\lambda_1 - \lambda_2) / \sum_{i=1}^{3} \lambda_i$) of the DTs at the source (*accepted*) and destination (*band*) voxels under consideration (Fig. 2.2). The C_L value is used as a reliability measure for the directions of the DT eigenvectors: the primary eigenvector is most reliable for high-C_L voxels ($C_L \geq 0.27$); the tertiary eigenvector for low-C_L voxels. If we denote the eigenvectors of the source voxel as $\mathbf{s}$, those of the destination voxel as $\mathbf{d}$, and the normal vector from source to destination as $\hat{\mathbf{n}}$, the speed function for the case in which both source and destination have high C_L (hC_L) is:

$$\underset{\mathrm{h}C_L \to \mathrm{h}C_L}{v} = \frac{1}{1 - \min\left(\left|\mathbf{s}_1 \cdot \mathbf{d}_1\right|^2, \left|\mathbf{s}_1 \cdot \hat{\mathbf{n}}\right|^2, \left|\mathbf{d}_1 \cdot \hat{\mathbf{n}}\right|^2\right)}$$

This is a simple metric based on the collinearity of the transition vector and the primary eigenvectors in the source and destination voxels. The relations for the remaining three cases can be found in [8]; all are defined similarly to the case given here—they differ only in weightings or the particular vectors considered as dictated by C_L. For Euclidean distance d between source and destination voxels, the arrival time for each *band* voxel is given by:

$$u\left(\mathbf{x}_{\mathrm{destination}}\right) = u\left(\mathbf{x}_{\mathrm{source}}\right) + d/v \tag{2}$$

Band voxels that lie in the neighborhoods of multiple *accepted* voxels retain only the lowest potential arrival time. Also, voxels not exceeding a minimum fractional anisotropy threshold (FA ≥ 0.2) are assigned $u\left(\mathbf{x}\right) = \infty$.

In contrast to [8], our implementation of FM additionally retains certain characteristics of front propagation as connectivity metrics. As each voxel is added to the *accepted* group, we store not only the arrival time for that voxel, but also the mean and minimum transition

velocities for front propagation along the path to that voxel. If γ represents the path of front propagation to the voxel centered at $\mathbf{x}$, then in addition to the arrival time $u(\mathbf{x})$, we also have:

$$v_{\text{mean}}(\mathbf{x}) = \left(\int_{\gamma} d\right)/u(\mathbf{x}) \quad (3)$$

$$v_{\min}(\mathbf{x}) = \min\left(v\left(\mathbf{x} \in \gamma\right)\right) \quad (4)$$

These latter measures provide inherent connectivity metrics derived directly from the tractography algorithm itself. The mean transition velocity at each voxel is a measure of the overall robustness of the connectional pathway from the seed to that point. In contrast, the minimum transition velocity along the front propagation path provides a measure of the rate-limiting step or a lower bound on the connection certainty. Notably, these connectivity measures can be refined, if desired, simply by redefining the speed function.

Several seed regions were manually labeled on the group mean FA image (results from one region are highlighted in Section 3). FM tractography was performed as described from each ROI for each subject and corresponding maps were generated for $u(\mathbf{x})$, $v_{\text{mean}}(\mathbf{x})$, and $v_{\min}(\mathbf{x})$. Processing time was less than one minute per ROI per subject on a 1.2 GHz machine.

2.3. Structural equation modeling

To determine the genetic and environmental contributions to these scalar connectivity measures, we rely on the accepted technique of structural equation modeling (SEM) in twin studies [14]. SEM evaluates contributions of additive genetic (A), shared environmental (C), and random environmental (E) components to the covariances of the observed variables (y) for MZ and DZ twins according to:

$$y_j = a\mathrm{A}_j + c\mathrm{C}_j + e\mathrm{E}_j \quad (5)$$

where $j \in \{1, 2\}$ indexes each twin in a pair (Fig. 3). In this work, our observed variables are the connectivity measures derived from FM tractography: arrival time, mean velocity, and minimum velocity as defined in Sec. 2.2. Since A, C, and E are unobservable (latent) variables, their weights $\theta = (a, c, e)$ are estimated by comparing the covariance matrix implied by the model ($\mathbf{\Sigma}(\theta)$) and the sample covariance matrix of observed variables ($\mathbf{S}$) using maximum-likelihood fitting:

$$\mathrm{F}_{\mathrm{ML},\theta} = \log|\mathbf{\Sigma}(\theta)| + \mathrm{Tr}\left(\mathbf{\Sigma}^{-1}(\theta)\,\mathbf{S}\right) - \log|\mathbf{S}| - p \quad (6)$$

where $p = 2$ is the number of observed variables. Under the null hypothesis that the population covariance matrix of the observed variables equals $\mathbf{\Sigma}(\theta)$, and the n-sample data are multivariate normal, $\mathrm{T}_{\mathrm{ML},\theta} = (n-1)\,\mathrm{F}_{\mathrm{ML},\theta}$ follows a χ^2 distribution with $p(p+1) - t$ degrees of freedom, where t is the number of free model parameters. Acceptance of the null hypothesis ($p > 0.05$) indicates a good fit for the model.

All statistical maps were further assessed using the false discovery rate (FDR) method to correct for multiple comparisons [15]. Statistical maps that could be thresholded in such a way that the $\mathrm{FDR} < 0.05$ were considered, by convention, to reach overall significance.

3. RESULTS

We present results from FM tractography and SEM for a seed region in the left superior longitudinal fasciculus (SLF)

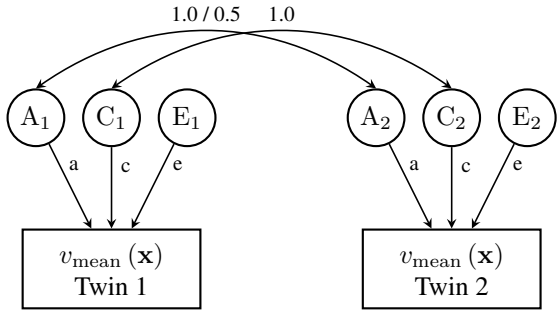

Fig. 2. Path diagram for the structural equation model for mean front propagation velocity. $v_{\text{mean}}(\mathbf{x})$ is determined by additive genetic (A), shared environmental (C), and unique environmental (E) factors. Contributions (straight arrows) of the A, C, and E factors to the observed mean velocity are weighted by the path coefficients a, c, and e respectively (assumed equal between co-twins). The correlation (curved arrow) between A1 and A2 is 1.0 for monozygotic pairs, 0.5 for dizygotic pairs. The correlation between C1 and C2 is always 1.0 by definition.

(Fig. 3). This major association bundle runs primarily in the anterior-posterior direction and contains fibers from the arcuate fasciculus which connect the language-associated regions of the left hemisphere: Broca's area in the inferior frontal gyrus and Wernicke's area in the superior temporal gyrus. The map of average arrival time across all subjects qualitatively confirms the reasonableness of the FM approach with the described speed function—we observe that the front evolves rapidly along directions of known connectivity, and very slowly towards unconnected regions like the insula. We used SEM to analyze maps of $v_{\text{mean}}(\mathbf{x})$ across the twin pairs to determine the genetic influence on this measure. The fraction of total variance in $v_{\text{mean}}(\mathbf{x})$ that was attributable to genetic sources (A, as opposed to common (C) or unique environmental (E) factors) is displayed as a the third map in Fig. 3. Finally, after statistical analysis and controlling for multiple comparisons, we derive a map illustrating brain regions whose connectivity to the seed, as quantified by $v_{\text{mean}}(\mathbf{x})$, is significantly ($p < 0.05$) influenced by genetic factors. We note that many of the voxels along the arcuate pathway reach the threshold for significance. This provides a possible structural basis for the conclusions of behavioral and functional studies linking genetic factors with verbal ability and IQ [16].

4. CONCLUSIONS

We developed a sequence of methods to study the heritability of neural connectivity using DWI. We introduced new scalar connectivity measures derived directly from a fast-marching tractography approach and we showed how, in a population of twins, structural equation modeling can decompose the variance in these metrics into genetic and environmental components. Finally, we have demonstrated the feasibility of this

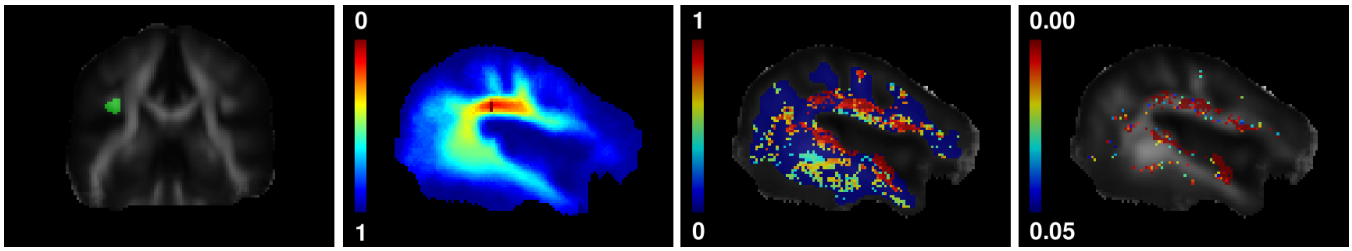

Fig. 3. Results of fast-marching tractography and structural equation modeling. Left: coronal view indicating placement of an ROI in the left superior longitudinal fasciculus, overlaid on the group average FA map. Mid-left: sagittal view of the group mean arrival time map (scaled as $\tanh(u(\mathbf{x}))$) at the level of ROI placement, illustrating the time-course of front evolution. Mid-right: sagittal view of the fraction of the variance of $v_{\text{mean}}(\mathbf{x})$ that was assigned by SEM to the additive genetic (A) component; *i.e.*, a normalized map of a^2 for $v_{\text{mean}}(\mathbf{x})$. Right: FDR-adjusted values for $p(\text{A})$, indicating the level of statistical significance of the genetic contributions to $v_{\text{mean}}(\mathbf{x})$; only voxels with $p < 0.05$ are shown; much of the arcuate fasciculus is highlighted.

approach by building, from a real data set, a map of brain regions whose connectivity to the SLF is significantly determined by genetic factors. In future work, we will refine the speed function to accommodate fiber orientation distribution functions that model multiple dominant directions per voxel. Ultimately, we plan to search, using genome-wide association, for specific genes that might modulate these connectivity metrics.

5. REFERENCES

[1] P. J. Basser, J. Mattiello, and D. LeBihan, "MR diffusion tensor spectroscopy and imaging," *Biophys. J.*, vol. 66, pp. 259–267, Jan 1994.

[2] V. J. Wedeen, P. Hagmann, W. Y. Tseng, T. G. Reese, and R. M. Weisskoff, "Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging," *Magn Reson Med*, vol. 54, pp. 1377–1386, Dec 2005.

[3] D. S. Tuch, "Q-ball imaging," *Magn Reson Med*, vol. 52, pp. 1358–1372, Dec 2004.

[4] V. Patel, Yonggang Shi, P. Thompson, and A. Toga, "Mesh-based spherical deconvolution for physically valid fiber orientation reconstruction from diffusion-weighted MRI," in *Biomedical Imaging: From Nano to Macro, 2009. ISBI '09. IEEE International Symposium on*, 28 2009-July 1 2009, pp. 614–617.

[5] A. Pfefferbaum, E. V. Sullivan, and D. Carmelli, "Genetic regulation of regional microstructure of the corpus callosum in late life," *Neuroreport*, vol. 12, pp. 1677–1681, Jun 2001.

[6] M. C. Chiang, M. Barysheva, D. W. Shattuck, A. D. Lee, S. K. Madsen, C. Avedissian, A. D. Klunder, A. W. Toga, K. L. McMahon, G. I. de Zubicaray, M. J. Wright, A. Srivastava, N. Balov, and P. M. Thompson, "Genetics of brain fiber architecture and intellectual performance," *J. Neurosci.*, vol. 29, pp. 2212–2224, Feb 2009.

[7] G.J.M. Parker, C.A.M. Wheeler-Kingshott, and G.J. Barker, "Estimating distributed anatomical connectivity using fast marching methods and diffusion tensor imaging," *Medical Imaging, IEEE Transactions on*, vol. 21, no. 5, pp. 505–512, May 2002.

[8] P. Staempfli, T. Jaermann, G. R. Crelier, S. Kollias, A. Valavanis, and P. Boesiger, "Resolving fiber crossing using advanced fast marching tractography based on diffusion tensor imaging," *Neuroimage*, vol. 30, pp. 110–120, Mar 2006.

[9] S. Jbabdi, P. Bellec, R. Toro, J. Daunizeau, M. Plgrini-Issac, and H. Benali, "Accurate anisotropic fast marching for diffusion-based geodesic tractography," *Int J Biomed Imaging*, vol. 2008, pp. 320195, 2008.

[10] T. E. Behrens, H. J. Berg, S. Jbabdi, M. F. Rushworth, and M. W. Woolrich, "Probabilistic diffusion tractography with multiple fibre orientations: What can we gain?," *Neuroimage*, vol. 34, pp. 144–155, Jan 2007.

[11] Y. Lu, A. Aldroubi, J. C. Gore, A. W. Anderson, and Z. Ding, "Improved fiber tractography with Bayesian tensor regularization," *Neuroimage*, vol. 31, pp. 1061–1074, Jul 2006.

[12] Ming-Chang Chiang, A.D. Leow, A.D. Klunder, R.A. Dutton, M. Barysheva, S.E. Rose, K.L. McMahon, G.I. de Zubicaray, A.W. Toga, and P.M. Thompson, "Fluid registration of diffusion tensor images using information theory," *Medical Imaging, IEEE Transactions on*, vol. 27, no. 4, pp. 442–456, April 2008.

[13] J. A. Sethian and A. Vladimirsky, "Ordered upwind methods for static Hamilton-Jacobi equations," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 20, pp. 11069–11074, 2001.

[14] J. E. Schmitt, R. K. Lenroot, S. E. Ordaz, G. L. Wallace, J. P. Lerch, A. C. Evans, E. C. Prom, K. S. Kendler, M. C. Neale, and J. N. Giedd, "Variance decomposition of MRI-based covariance maps using genetically informative samples and structural equation modeling," *Neuroimage*, vol. 47, pp. 56–64, Aug 2009.

[15] J. D. Storey and R. Tibshirani, "Statistical significance for genomewide studies," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 100, pp. 9440–9445, Aug 2003.

[16] Rosa A. Hoekstra, Meike Bartels, and Dorret I. Boomsma, "Longitudinal genetic study of verbal and nonverbal IQ from early childhood to young adulthood," *Learning and Individual Differences*, vol. 17, no. 2, pp. 97 – 114, 2007.