Supplementary materials

I. *NTU-DSI-122 DSI template*

We previously developed a DSI template, NTU-DSI-122, in the standard ICBM152 space (available at https://www.nitrc.org/projects/ntu-dsi-122). This template was built through a novel two-step registration strategy, which incorporated both the macroscopic anatomical information using high-resolution T1W images and the microscopic structural information obtained from DSI datasets. Briefly, in the first step, an intermediate DSI template was constructed in the ICBM152 space by averaging the temporarily registered DSI datasets using deformation maps estimated through T1W images; in the second step, the temporarily registered DSI datasets were further registered to the intermediate DSI template using the LDDMM-DSI algorithm (Hsu et al., 2012), yielding the final registered DSI datasets. The NTU-DSI-122 template was constructed by averaging the final registered DSI datasets. Therefore, one can consider the first step as a transformation emphasizing on conforming brain morphology to the ICBM152 space, and the second step as a transformation emphasizing the q-space signals convergent across subjects (Hsu et al., 2014). By using this two-step registration strategy, NTU-DSI-122 showed high anatomical consistency in matching to the standard ICBM152 template. The final template was the average of 122 registered DSI datasets from healthy adults. In addition to the high

signal-to-noise ratio of the template, it had several unique qualities. (1) All the datasets were acquired on a 3T MRI machine with a 32 channel phased arrayed head coil at isotropic spatial resolution of 2.5 x 2.5 x 2.5 mm^3. (2) All the datasets passed the quality assurance to ensure that the head motion-induced signal dropout was lower than 3 images out of 5508 images in total for each subject. (3) Susceptibility-induced distortion was corrected for all the datasets using the field maps (Hsu et al., 2009). (4) An advanced two-step registration procedure was used to ensure high consistency in both brain anatomy and q-space signals.

II. *Tractography procedures*

The voxels with GFA values greater than a given threshold of 0.05, which was two standard deviations above the mean GFA values of the gray matter, were selected as the white matter regions and were used as seed voxels for tractography over the whole brain. Consistent with previous reports, the GFA value derived from HARDI data was mostly linear with but lower than the FA value derived from DTI (Fritzsche et al., 2010; Gorczewski et al., 2009), and our threshold value of 0.05 for white matter masking was comparable with the GFA value of 0.05 used by (Berman et al., 2008) based on HARDI data. For each seed voxel, the proceeding orientation for the next step was decided by the angular deviation between the primary orientation within the seed voxel and all of the fiber orientations of its nearest voxels; the most coincident orientation with the minimum angular deviation was chosen. The range of the angle thresholds for the 76 long tract bundles was from 28° to 40°. The new starting point was obtained by moving the seed point with a proceeding length of one voxel (1 mm) for each step along the most coincident orientation. When the angle deviations were greater than a given angular threshold, the tracking would stop. To ensure the accuracy of the tractography, only the tracts with lengths falling between 37.5 mm and 173.8 mm (mean: 70.4 mm; median: 62.5 mm) were selected, and a range of tract numbers from 100 to 500 was required for each tract bundle.

III. *Threshold-Free Cluster Enhancement*

Because the GFA values sampled along a tract bundle are derived from a continuous function, the magnitude of differences between groups in each measured step should have spatial contiguity along a tract. We used TFCE to improve the statistical power and exploited spatially extended differences. TFCE reweighted each data point (e.g., voxel) of a statistical image with its own magnitude and neighboring supportive segments (Smith and Nichols, 2009). The TFCE output at data point p was

$$TFCE(p) = \int_{h=h_0}^{h_p} e(h)^E h^H dh$$

where h_p was the statistical magnitude (e.g., t statistic) at a given point p. As h varied between height zero, h_0, to h_p, the cluster extent that survived at the given threshold h was e. The predetermined parameters, E and H, assigned weights to the cluster extent and statistical magnitude, respectively. TFCE was the sum of the products from a height of zero, h_0, to h_p.

In the original paper of TFCE (Smith and Nichols, 2009), H = 2 was suggested for the t-statistic image and E = 0.5 was suggested for the three dimensional data. In our TBAA method, the dimensionality of statistical images was reduced to 1D. Therefore, the parameter E had to be modified. Based on our simulation for 1D data,

the parameter E was chosen to be 2.

Controlling for the FWER

A permutation test based on a TFCE-weighted statistical image was used for statistical inference (Smith and Nichols, 2009). A maximal statistic approach was used to control the FWER (Nichols and Hayasaka, 2003). For each permutation, subjects were randomly assigned into two groups and the largest TFCE weighted statistic among the images was derived. After 5000 permutations, the results approximated a cumulative density function of maximal TFCE statistic, which was used to calculate the FWER adjusted p values.