scientific **data**

OPEN

DATA DESCRIPTOR

A diffusion MRI tractography atlas for concurrent white matter mapping across Eastern and Western populations

Yijie Li[1,7], Wei Zhang[1,7], Ye Wu[2], Li Yin[3], Ce Zhu[1], Yuqian Chen[4], Suheyla Cetin-Karayumak[5], Kang Ik K. Cho[5], Leo R. Zekelman[4], Jarrett Rushmore[5,6], Yogesh Rathi[5], Nikos Makris[5], Lauren J. O'Donnell[4 ✉] & Fan Zhang[1 ✉]

The study of brain differences across Eastern and Western populations provides vital insights for understanding potential cultural and genetic influences on cognition and mental health. Diffusion MRI (dMRI) tractography is an important tool in assessing white matter (WM) connectivity and brain tissue microstructure across different populations. However, a comprehensive investigation into WM fiber tracts between Eastern and Western populations is challenged due to the lack of a cross-population WM atlas and the large site-specific variability of dMRI data. This study presents a dMRI tractography atlas, namely the *East-West WM Atlas*, for concurrent WM mapping between Eastern and Western populations and creates a large, harmonized dMRI dataset (n=306) based on the Human Connectome Project and the Chinese Human Connectome Project. The curated WM atlas, as well as subject-specific data including the harmonized dMRI data, the whole brain tractography data, and parcellated WM fiber tracts and their diffusion measures, are publicly released. This resource is a valuable addition to facilitating the exploration of brain commonalities and differences across diverse cultural backgrounds.

Background & Summary

Differences in genetics, cultural backgrounds, and environmental influences play a vital role in shaping the structure and function of the human brain[1,2]. There is a surge of interest in studying the brains of Eastern and Western populations to investigate potential cultural influences on cognition and mental health[2–8]. Noninvasive neuroimaging is an important tool to uncover the neural basis of the human mind and behavior. In particular, magnetic resonance imaging (MRI) has greatly improved the understanding of cross-cultural brain differences in morphology, cortical thickness, and other aspects between Eastern and Western populations[6–14].

Diffusion MRI (dMRI) is an advanced MRI technique that can probe the diffusion of water molecules in biological tissues to characterize the underlying microstructure[15]. In particular, dMRI enables a computational process, namely tractography, that uniquely enables in-vivo reconstruction of the brain's white matter (WM) connections at macro scale[16]. Quantitative approaches using tractography have become popular tools for studying the brain's connectivity and tissue microstructure, including connectome-based analyses to investigate the structural connectivity of the entire brain and tract-specific analyses to investigate particular anatomical fiber tracts[17]. Currently, dMRI tractography plays a prominent role in studying the brain's WM connections in health and disease[17–20].

Multiple studies have utilized dMRI tractography to explore cross-cultural brain studies between Eastern and Western populations. Existing studies focus on connectome-based analyses[2,21,22], which have revealed distinct brain topographic features specific to each population in terms of whole brain structural connectivity strengths.

[1]School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China. [2]School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China. [3]West China Hospital of Medical Science, Sichuan University, Chengdu, China. [4]Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, USA. [5]Department of Psychiatry, Brigham and Women's Hospital, Harvard Medical School, Boston, USA. [6]Department of Anatomy and Neurobiology, Boston University School of Medicine, Boston, USA. [7]These authors contributed equally: Yijie Li, Wei Zhang. ✉e-mail: odonnell@bwh.harvard.edu; fan.zhang@uestc.edu.cn

However, investigation of specific anatomical fiber tracts (e.g., the arcuate fasciculus and the corticospinal tract) between Eastern and Western populations is lacking. Unlike connectome-based analyses that are data-driven to explore the structural connectivity of the entire brain, tract-specific analyses enable hypothesis-driven studies of certain anatomical fiber tracts[17]. This allows detailed investigation of local WM regions in association with certain brain functions, an approach that has been widely used in neuroscientific and clinical research studies[23–27]. Yet, a comprehensive investigation into WM fiber tracts between Eastern and Western populations is still missing due to the following challenges.

The first challenge is the lack of dMRI tractography atlases that enable concurrent mapping of WM fiber tracts between Eastern and Western populations. In neuroimaging research, brain atlases serve as essential tools for standardized representation of the anatomical structures and functional regions within the brain[28–30]. Creating a brain atlas generally involves constructing a common template derived from a group of individuals, which can be applied to diverse individuals to identify subject-specific anatomically segregated brain regions. Previous studies examining the brains across different cultures, focusing on the cortical surface and brain morphology[10,31,32], have highlighted the necessity of developing atlases specific to each population. This is crucial to accurately capture subtle anatomical variations that exist within a particular cultural group. In the dMRI tractography literature, existing WM atlases primarily originate from data collected from Western populations[33–38]. While these atlases can be directly applied to data from an Eastern population, they might not capture the subtle inter-population anatomical variability of the WM. Therefore, there is a high demand for an across-population WM atlas that facilitates concurrent mapping of WM fiber tracts between Eastern and Western populations, promoting more inclusive cross-culture brain studies.

A second challenge in studying cross-cultural brain WM differences between Eastern and Western populations is the site-specific variability in dMRI data. Unlike other types of group-wise analyses (e.g., health vs disease) where neuroimaging data of the participants can be acquired on the same MRI scanner from a single acquisition site, cross-cultural analyses usually rely on data acquired at multiple sites from different countries or continents. In dMRI, it is well known that variations in site-specific data arising from different MRI scanners and/or acquisition protocols can result in significant biases in measures of WM connectivity and microstructure[39–45]. Thus, many studies underscore the significance of eliminating site-specific biases before running joint analysis in multi-site studies[40,41,46–52]. dMRI harmonization is an effective way to mitigate measurement differences attributed to scanner-, protocol-, or other site-related effects, and it has been used in recent studies to enable pooled dMRI data analyses[40,41,46–51]. However, there are currently no harmonized dMRI resources to enable joint dMRI analysis between Eastern and Western populations.

In light of the above, this study presents a novel dMRI tractography atlas for concurrent WM mapping across Eastern and Western populations, referred to as the *East-West WM Atlas*. High-quality dMRI data from the Human Connectome Project (HCP)[53] and the Chinese Human Connectome Project (CHCP)[2], acquired from Western and Eastern populations, respectively, are used for the atlas creation. To remove site-specific biases from the multi-site data, we employ our advanced dMRI harmonization algorithm[41,46] that reconciles the raw dMRI signals across the two datasets while preserving inter-population variability. This generates a pooled large-scale dMRI dataset including a total of 306 individuals (153 for each cultural population). From this dataset, we create a cross-population WM atlas using our well-established fiber clustering pipeline[36,54,55] that enables simultaneous tractography atlas creation across populations. This generates an anatomically curated atlas including a tract-level parcellation of major anatomical fiber tracts as well as a fine-scale parcellation of the entire WM. Importantly, the created atlas, together with the harmonized datasets, enables subject-specific WM parcellation for direct comparison of the WM connectivity and microstructural measures between the two populations. Constituting an essential new resource for cross-cultural brain studies, all code of related computational tools and the data including the harmonized dMRI dataset, the created WM atlas, the subject-specific WM tracts, and the associated diffusion measures are openly available.

Methods

This section introduces the dMRI dataset and the computational steps for developing the proposed WM atlas (See Fig 1 for a method overview).

Datasets and Participants.

This study utilized dMRI datasets from two sources: the Human Connectome Project Young Adult (HCP-YA) database, obtained from participants living in a Western culture from the USA[53], and the Chinese Connectome Project (CHCP) database, obtained from participants living in an Eastern culture from China[2]. The HCP-YA database comprised dMRI data from over 1,000 young healthy adults (22 to 37 years old), while the CHCP database included dMRI data from over 300 healthy adults including teenagers, young adults, and elderly adults (18 to 79 years old). In our investigation, we used two subsets of participants from CHCP and HCP-YA. We included the young adult participants (22 to 37 years) from the CHCP dataset and selected a subset of HCP-YA participants with similar age and sex distributions. Overall, in our study, each dataset contained data from 153 participants (HCP-YA: 24.2 ± 1.4 years, 72 females and 81 males; CHCP: 23.9 ± 2.4 years, 68 females and 85 males; no significant age or sex differences), resulting in a large cohort of 306 participants used for the atlas creation and technical validation.

MRI Acquisitions and Preprocessing.

The dMRI data of the above participants were publicly available in the HCP-YA (db.humanconnectome.org) and CHCP (https://doi.org/10.11922/sciencedb.01374) databases. The detailed acquisition parameters are described in[53] and[2]. Briefly, the dMRI data in HCP-YA was acquired on a customized 3T Connectome Siemens Skyra scanner, with the following acquisition parameters: $TE = 89.5\,ms$, $TR = 5520\,ms$, voxel size $= 1.25 \times 1.25 \times 1.25\,mm^3$, and a total of 288 volumes including 18 baseline images and 90 diffusion-weighted images at each of the three shells of $b = 1000/2000/3000\,s/mm^2$. The dMRI data in CHCP
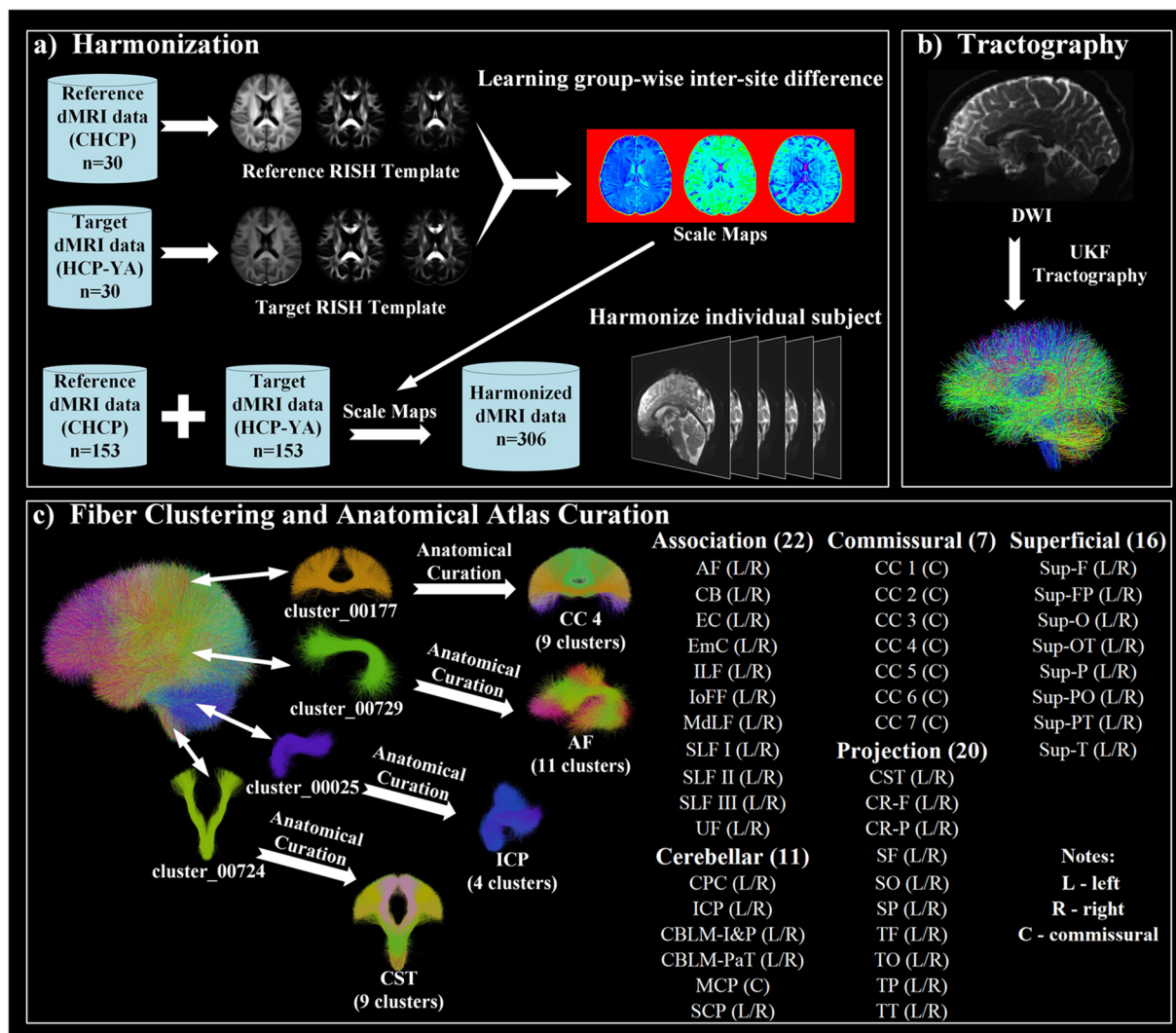
Fig. 1 Method overview: (**a**) dMRI data harmonization to eliminate inter-site discrepancies between the CHCP and HCP datasets. (**b**) Whole brain tractography for the reconstruction of WM connections from the entire brain. (**c**) Group-wise fiber clustering atlas creation concurrently using tractography data from the CHCP and HCP populations. A total of 76 anatomical fiber tracts categorized into 5 tract categories are included.

was acquired on a 3T Siemens Prisma MRI scanner with the following acquisition parameters: $TE = 86\,ms$, $TR = 3500\,ms$, voxel size $= 1.5 \times 1.5 \times 1.5\,mm^3$, 14 baseline images at b = 0, 93 diffusion-weighted images at $b = 1000\,s/mm^2$, and 92 diffusion-weighted images at $b = 2000\,s/mm^2$. In both dMRI datasets, we used only single-shell dMRI data for reasonable computation time and memory use when performing tractography. We employed the $b = 2000\,s/mm^2$ data as these were available in both databases, and the angular resolution is better and more accurate at higher b-values[56,57].

Both the HCP-YA and CHCP datasets were preprocessed with standard dMRI processing steps, including brain masking, eddy current-induced distortion correction, motion correction, and EPI distortion correction. Specifically, for the HCP-YA data, we used the dMRI data already preprocessed with the HCP minimum processing pipeline[58]. For the CHCP data, we used our well-established dMRI data processing pipeline[36] (https://github.com/pnlbwh/pnlpipe/tree/v2.2.0), including: CNN brain masking[59] (https://github.com/pnlbwh/CNN-Diffusion-MRIBrain-Segmentation/tree/v0.3) for dMRI images, and eddy current, motion, and EPI distortion correction using FSL topup and eddy tools (version 6.0.6.5)[60].

Data Harmonization.   While the HCP-YA and CHCP dMRI datasets underwent similar preprocessing, there were inherent scanner-specific biases due to the involvement of multiple acquisition sites, scanners, and acquisition protocols. To mitigate the inter-site variability within this multi-site dMRI study, we employed our retrospective harmonization algorithm[41,46]. Our harmonization method relied on rotation-invariant spherical harmonics (RISH) features at the dMRI signal level. Its primary objective was to remove scanner-specific biases across datasets, accounting for the non-linearities in dMRI data, which may vary by region and tissue. This method was successfully used in multiple studies for dMRI harmonization to enable pooled data analysis[59,61–63].
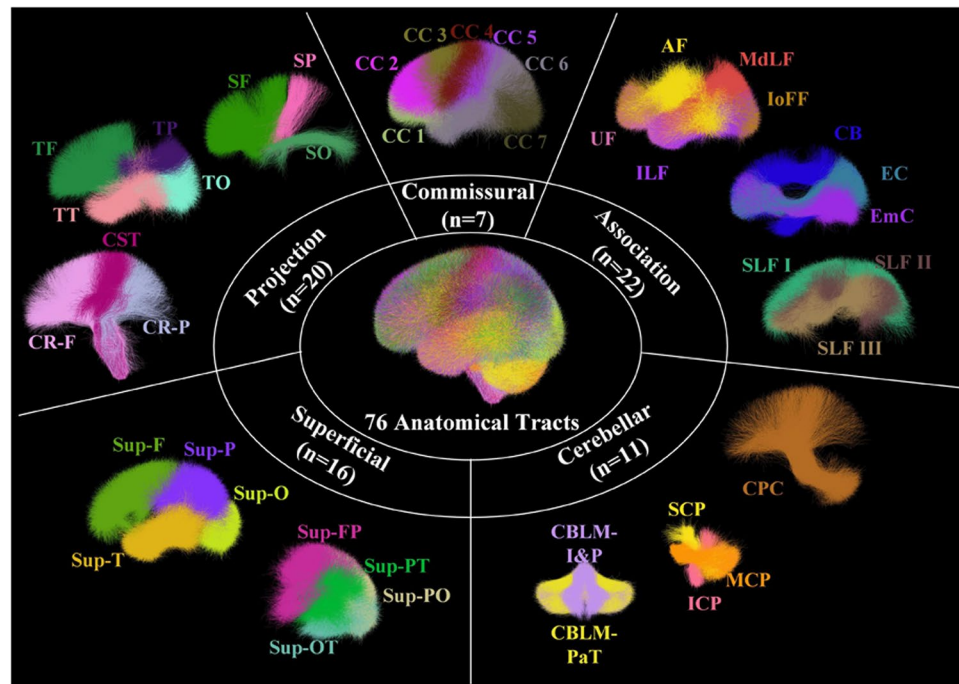
Fig. 2 Visualization of anatomical fiber tracts included in the proposed atlas, categorized into five tract categories. Association tracts include: arcuate fasciculus (AF), cingulum bundle (CB), external capsule (EC), extreme capsule (EmC), inferior longitudinal fasciculus (ILF), inferior occipito-frontal fasciculus (IoFF), middle longitudinal fasciculus (MdLF), superior longitudinal fascicle I, II, III (SLF I, II, III), and uncinate fasciculus (UF). Cerebellar tracts include: cortical-ponto-cerebellar (CPC), inferior cerebellar peduncle (ICP), intracerebellar input and Purkinje tract (CBLM-I&P), intracerebellar parallel tract (CBLM-PaT), middle cerebellar peduncle (MCP), and superior cerebellar peduncle (SCP). Projection tracts include: corticospinal tract (CST), corona-radiata-frontal (CR-F), corona-radiata-parietal (CR-P), striato-frontal (SF), striato-occipital (SO), striato-parietal (SP), thalamo-frontal (TF), thalamo-occipital, thalamo-parietal (TP), and thalamo-temporal (TT). Commissural tracts include seven segments of the corpus callosum (CC): rostrum (CC 1), genu (CC 2), rostral body (CC 3), anterior midbody (CC 4), posterior midbody (CC 5), isthmus (CC 6), and splenium (CC 7). Superficial tracts include: superficial-frontal (Sup-F), superficial-frontal-parietal (Sup-FP), superficial-occipital (Sup-O), superficial-occipital-temporal (Sup-OT), superficial-parietal (Sup-P), superficial-parietal-occipital (Sup-PO), superficial-parietal-temporal (Sup-PT), and superficial-temporal (Sup-T).

The harmonization process (https://github.com/pnlbwh/dMRIharmonization/tree/v2.1) comprised the following two essential steps. Briefly, the first step was to learn dMRI differences between the two datasets while preserving inter-subject biological variability at a group level. To do so, we selected 30 subjects from each of the CHCP and HCP datasets with similar age and sex distributions (no significant differences). For each selected subject, we calculated RISH features to capture orientation-independent microstructural tissue properties and enable the reconstruction of the harmonized dMRI signal[64]. Nonlinear mappings of the RISH features were learned from the HCP-YA (target site) to the CHCP data (reference site). While CHCP is selected as a reference site, we have previously shown that the choice of the reference site does not affect the performance of harmonization[41]. The second step of the dMRI harmonization process was to harmonize the dMRI data in each individual subject. This was done by applying the learned mappings to the RISH features computed from the subjects in the target site. This generated harmonized RISH features that correspond to the reference site. Then, from the harmonized RISH features, dMRI signals were reconstructed for the following tractography analyses.

Whole Brain Tractography. Whole brain tractography was performed using our two-tensor Unscented Kalman Filter (UKF) approach[65,66], as implemented in the ukftractography package (https://github.com/pnlbwh/ukftractography). The UKF method fitted a mixture model of two tensors to the dMRI data while tracking fibers, employing prior information from the previous step to help stabilize model fitting. UKF was shown to be highly consistent for fiber tracking in dMRI data from independently acquired populations across ages, health conditions, and dMRI acquisitions[36,67,68].

For each of the harmonized dMRI scans from the HCP-YA and CHCP datasets, we adopted the tractography parameters as used in[36] for whole brain WM fiber tracking. In brief, tractography was seeded in all voxels within the brain mask where fractional anisotropy (FA) was greater than 0.1. Tracking stopped where the FA value fell below 0.08 or the normalized mean signal fell below 0.06. To ensure comparability and minimize bias related to variations in streamline counts across subjects and datasets, we uniformly downsampled the whole-brain tractography for each subject to 600,000 streamlines. The diffusion tensors calculated during tractography, together
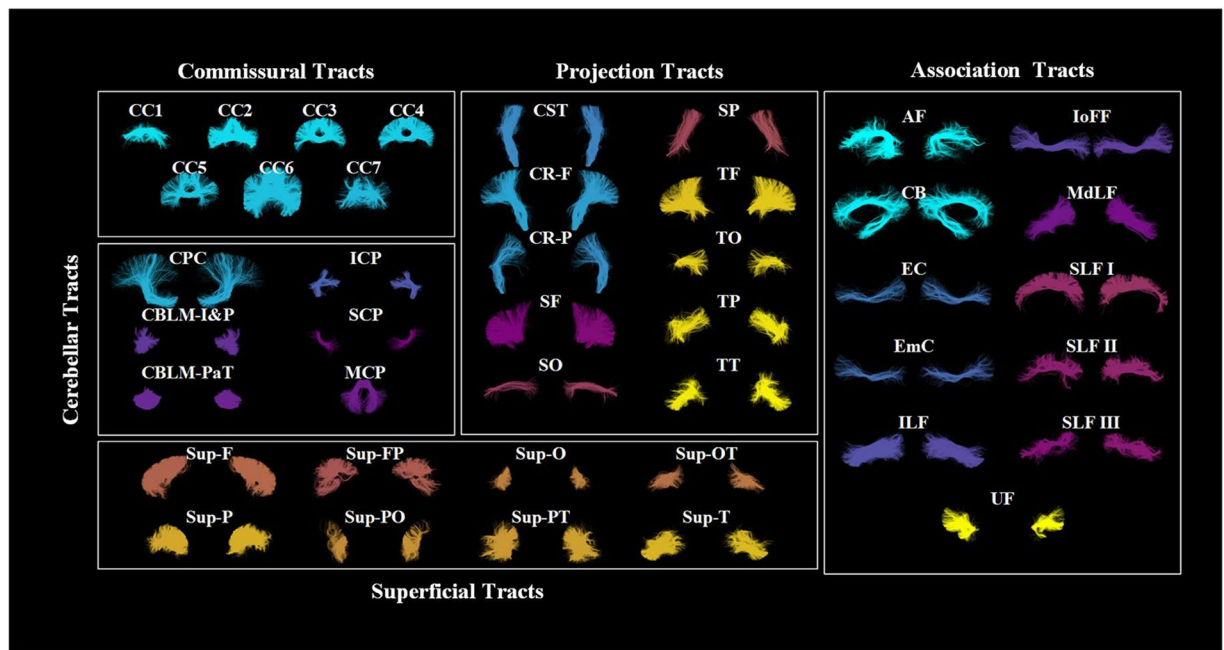
Fig. 3 Visualization of the anatomical fiber tracts identified in one example subject.

with their FA and mean diffusivity (MD) values, were preserved for subsequent analyses to enable a comprehensive assessment of the WM microstructure.

Fiber Clustering Atlas Creation. After obtaining the tractography data per subject, a groupwise whole-brain fiber clustering atlas was created using our robust, data-driven fiber clustering pipeline[36,54,55], as implemented in the *whitematteranalysis* (WMA) software (https://github.com/SlicerDMRI/whitematteranalysis). The pipeline included two key processes: a group-wise tractography registration to align the tractography of all subjects to a common space, and a spectral clustering of tractography to subdivide the registered tractography data into multiple fiber clusters simultaneously. The WMA fiber clustering pipeline has been successfully used for creating WM[36] and cranial nerve[69,70] tractography atlases.

In our study, we pooled the tractography data of the 306 subjects from the CHCP and HCP datasets together for atlas creation. This allowed us to learn a tractography atlas for concurrent mapping of the WM connections across the two populations. Specifically, from each subject's whole brain tractography, 10,000 streamlines were randomly selected, resulting in approximately 3 million streamlines for clustering. The WMA fiber clustering pipeline was used to parcellate all streamlines into K clusters. We chose $K = 800$, which was shown to be a good parcellation scale of the whole brain WM in previous studies[36,71,72]. We then performed anatomical curation of the fiber clusters by annotating each fiber cluster with an anatomical label belonging to a certain anatomical tract (e.g., the corticospinal tract) or an unclassified category. To do so, we leveraged the ORG atlas built only using the HCP data as a reference[36,73]. We performed co-registration of the two atlases using a tractography-based registration[54] and calculated the mean closest point distances[55,74] between the clusters in the new atlas and those in the ORG atlas and then assigned each new atlas cluster with the label of the closest ORG cluster. In total, the proposed atlas annotates 60 deep WM tracts and 222 superficial fiber clusters categorized into 16 groups based on their associated brain lobes, as detailed in Fig. 2.

WM Parcellation of Individual Subjects. With the created WM atlas, we parcellated the whole brain tractography data of each subject in the HCP-YA and CHCP datasets, using the subject-specific fiber clustering tool provided in WMA. In brief, subject-specific tractography data was spectrally embedded into the atlas space by computing the spatial distances of its streamlines to the atlas, followed by the assignment of each fiber to the closest atlas cluster. As a result, the new subject's tractography was divided into multiple fiber clusters, where each cluster corresponded to a certain atlas fiber cluster. Anatomical tract identification of each subject was conducted by finding the subject-specific clusters that corresponded to the annotated atlas clusters. Figure 3 displays a subject-specific parcellation, demonstrating all anatomical fiber tracts. Then, dMRI measures were extracted from each parcellated cluster and anatomical tract using SlicerDMRI[75,76] for all subjects. These included the widely used dMRI measures: fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AD), radial diffusivity (RD), number of streamlines (NoS), number of streamline points (NoP) and streamline length.

Data Records

Our neuroimaging dataset is now accessible on Science Data Bank[77], structured following the Brain Imaging Data Structure (BIDS) format[78]. In the root folder, demographic information of the CHCP and HCP-YA cohorts is provided in the "participants.tsv" file, and the harmonized dMRI data and the brain mask of each subject under study are provided in Neuroimaging Informatics Technology Initiative (NIfTI) format. In the
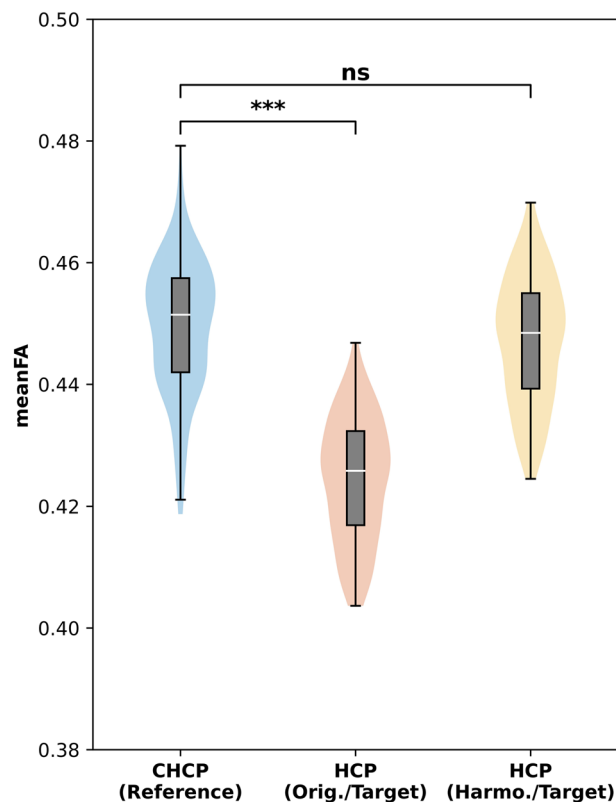
Fig. 4 Effects of dMRI harmonization in terms of the mean FA of the entire brain's WM skeleton. Notes: "*ns*" stands for no significance where p-value > 0.05, "*" stands for p-value < 0.05, "**" stands for p-value < 0.01, and "***" stands for p-value < 0.001.

'derivatives' directory, there are two subfolders corresponding to the created WM tractography atlas and the subject-specific WM parcellation results. Specifically, the "derivatives/atlas" folder provides the group-wise fiber clustering atlas including 800 clusters stored in VTK format, the anatomical tract label annotation organized in a 3D Slicer[79] scene file (in medical reality modeling language - MRML, an XML format), and the population mean b0 image by transforming all subjects' b0 images into the atlas space stored in Neuroimaging Informatics Technology Initiative (NIfTI) format. Instructions about how to visualize the atlas using the 3D Slicer software via SlicerDMRI are provided in a "README.md". The "derivatives/tractography" provides the whole-brain tractography data, the parcellated fiber clusters, and the anatomical fiber tracts, which are all stored in VTK format. The dMRI measures of each fiber tract (such as FA, MD, and NoS) are also provided in CSV format.

Technical Validation

This section focuses on demonstrating the effectiveness of the dMRI harmonization process, the accuracy of the atlas, and its consistency across diverse populations. Additionally, a preliminary comparison of the CHCP and HCP datasets using the newly proposed atlas is conducted, providing further insights into the applicability of the atlas.

Effects of dMRI Harmonization to Remove Site-Specific Biases. We first examined the impact of the harmonization procedure. This was done by comparing the voxel-wise diffusion measure FA computed from the dMRI data before and after harmonization at the target site (HCP-YA), in comparison to the reference site (CHCP). The comparison was performed at two different scales, including the entire brain's WM skeleton and 42 WM regions of interest (ROIs), as delineated by the Illinois Institute of Technology's (IIT) atlas[80,81]. All 306 subjects from the CHCP and HCP-YA datasets under study were included in the analysis. The mean FA values between the reference and the original target data and between the original and the harmonized target data were compared. First, independent t-tests were performed to compare the entire brain's WM skeleton. Next, we performed an ROI-based comparison to assess harmonization's performance in different regions of the brain. Independent t-tests were performed for each ROI, followed by Bonferroni corrections performed across the multiple comparisons across all ROIs.

Figure 4 gives the mean FA over the entire WM skeleton across the CHCP dataset, the original HCP-YA dataset, and the harmonized HCP-YA dataset. We can observe that compared to the original HCP-YA dataset,
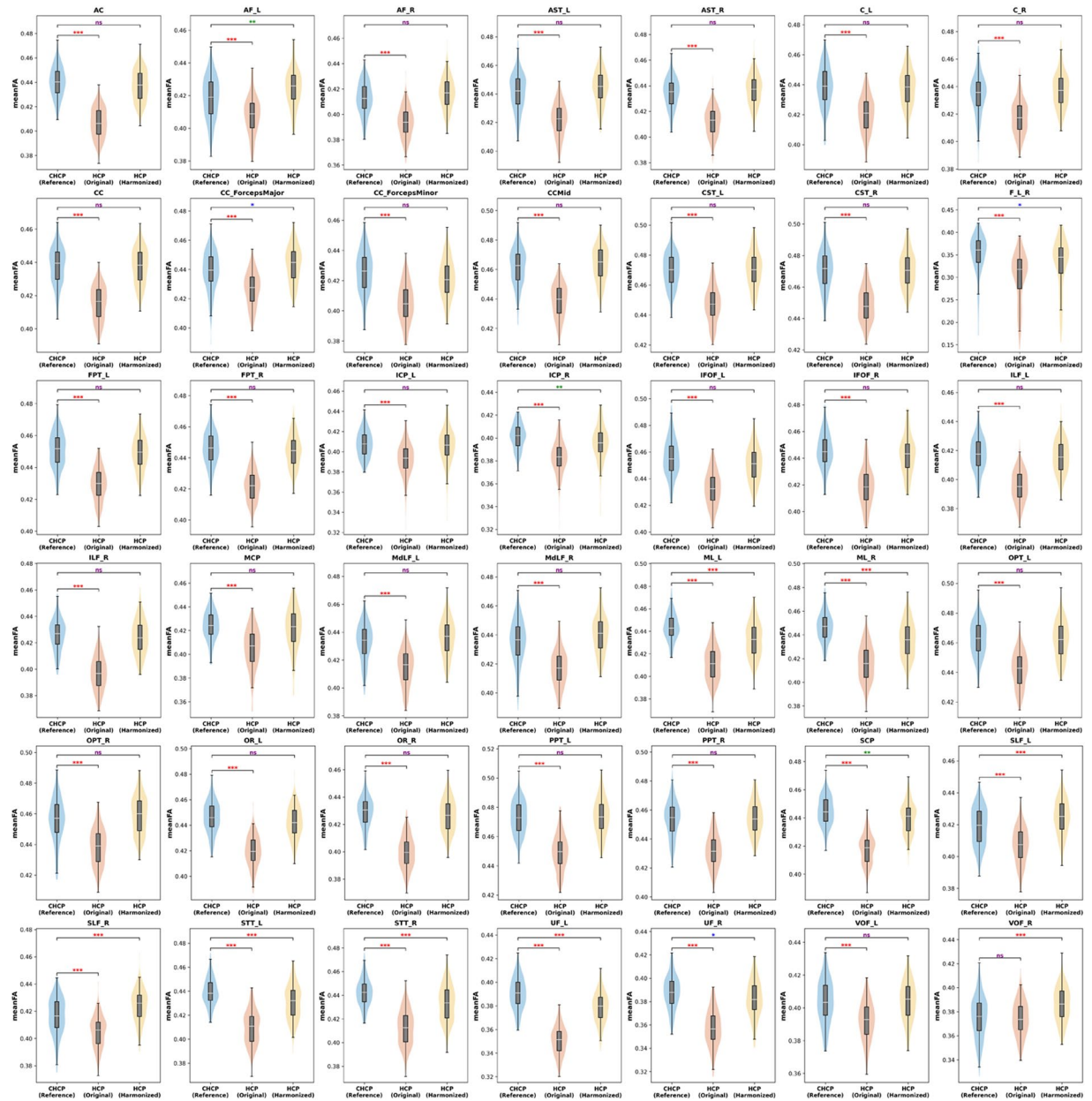
Fig. 5 Effects of dMRI harmonization on WM ROI analyses. Notes: purple "*ns*" stands for no significance where p-value > 0.05, blue "*" stands for p-value < 0.05, green "**" stands for p-value < 0.01, and red "***" stands for p-value < 0.001.

the FA values of the harmonized HCP-YA dataset became more similar to the CHCP dataset. Notably, the significant differences between the CHCP data and the original HCP-YA data ($p \ll 0.01$) were reduced and statistically diminished ($p = 0.06$) after harmonization. Figure 5 gives the comparison results on the 42 ROIs for a more detailed assessment. After harmonization, in all ROIs, the FA values of the harmonized data were closer to those of the reference data, where 14 out of the 42 ROIs showed no significant differences anymore. These indicated that the harmonization process effectively reduced the discrepancies while preserving potential true anatomical variability due to genetic or cultural differences. Overall, the above results underscored the effectiveness of harmonization in eliminating scanner-related differences between the original and reference datasets.

Anatomical Plausibility of Curated Fiber Tracts. We then assessed the anatomical plausibility of our proposed atlas. This was done by comparing our proposed atlas to the existing ORG atlas[36], which has been previously used in many neuroscientific and clinical research studies[26,59,82–86]. The evaluation was performed by measuring the spatial similarity of corresponding curated fiber tracts in both atlases using the weighted Dice (wDice) coefficient[73,87] - a metric designed for measuring fiber tract spatial overlap. A wDice score over 0.72 is suggested to be a good tract overlap[73,87].
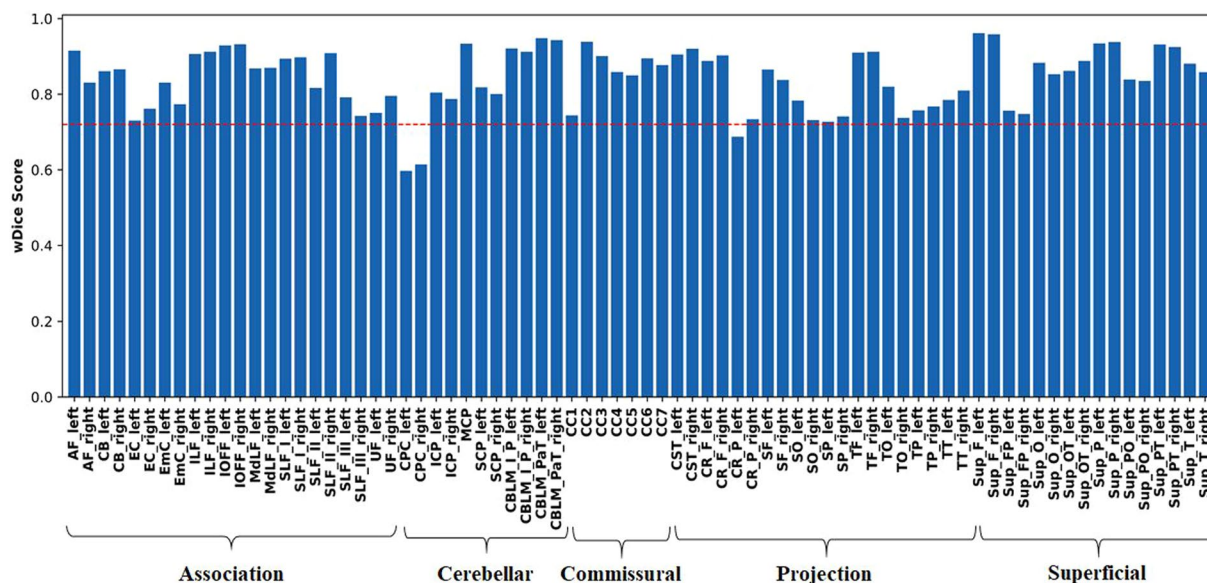
Fig. 6 wDice score for each of the corresponding anatomical fiber tracts in the ORG atlas and the proposed atlas.
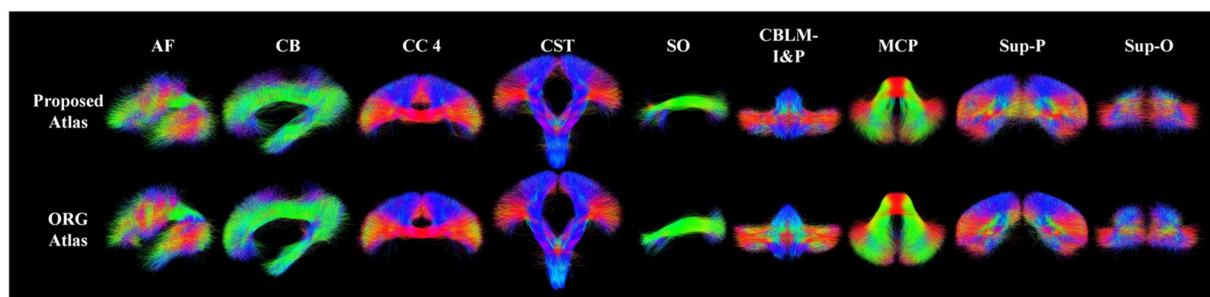

Fig. 7 Visual comparison of example tracts of the proposed atlas and the ORG atlas.

Figure 6 gives the wDice score for each fiber tract. On average, the mean wDice score across all fiber tracts was 0.84, showing a high volumetric consistency with the ORG atlas. Notably, 73 out of the 76 tracts have a wDice score over 0.72. Figure 7 gives a visual comparison of example tracts from the two atlases, where we can observe highly visually comparable tracts between the two atlases.

Enabling Concurrent WM Mapping Across Populations. We evaluated the ability of the proposed atlas to enable concurrent mapping of WM tracts in the CHCP and HCP-YA datasets. This was done by comparing it to the ORG atlas that was created purely using the HCP data. To do so, the whole brain tractography of each of the CHCP and HCP-YA subjects was parcellated using the proposed atlas and the ORG atlas separately. Three quantitative metrics were computed to evaluate the WM parcellation performance. The first metric was the WM parcellation generalization (WMPG)[36,88], which was measured as the percentage of successfully identified fiber clusters across all subjects in each dataset. A high WMPG value signified strong generalizability of the atlas to a given subject, reflecting the applicability of the atlas across diverse brain structures. The second metric was the inter-subject parcellation variability (ISPV)[36,89], which was measured as the coefficient of variation (CoV) of NoS for each fiber cluster across each population. A low ISPV value indicated low variability in parcellation, thus signifying a high consistency of WM parcellation across subjects. The third metric was tissue microstructure homogeneity (TMH), which was measured as the CoV of the mean cluster FA values across the subjects in each population[90]. A lower TMH indicates a higher homogeneity of tract microstructural properties as a proxy for better parcellation.

Figure 8 presents the quantitative comparison results between employing our proposed atlas and the ORG atlas for subject-specific white matter parcellation. For each metric and each population, an independent t-test was performed to compare the measures between the two atlases. Significant improvement can be observed in the CHCP dataset using the proposed atlas, as indicated by the increased WMPG and decreased ISPV metrics ($p \ll 0.01$). We can also see a lower TMH using the proposed atlas with a very low p-value ($= 0.058$) though without statistical significance. These enhancements demonstrated the effectiveness of incorporating CHCP
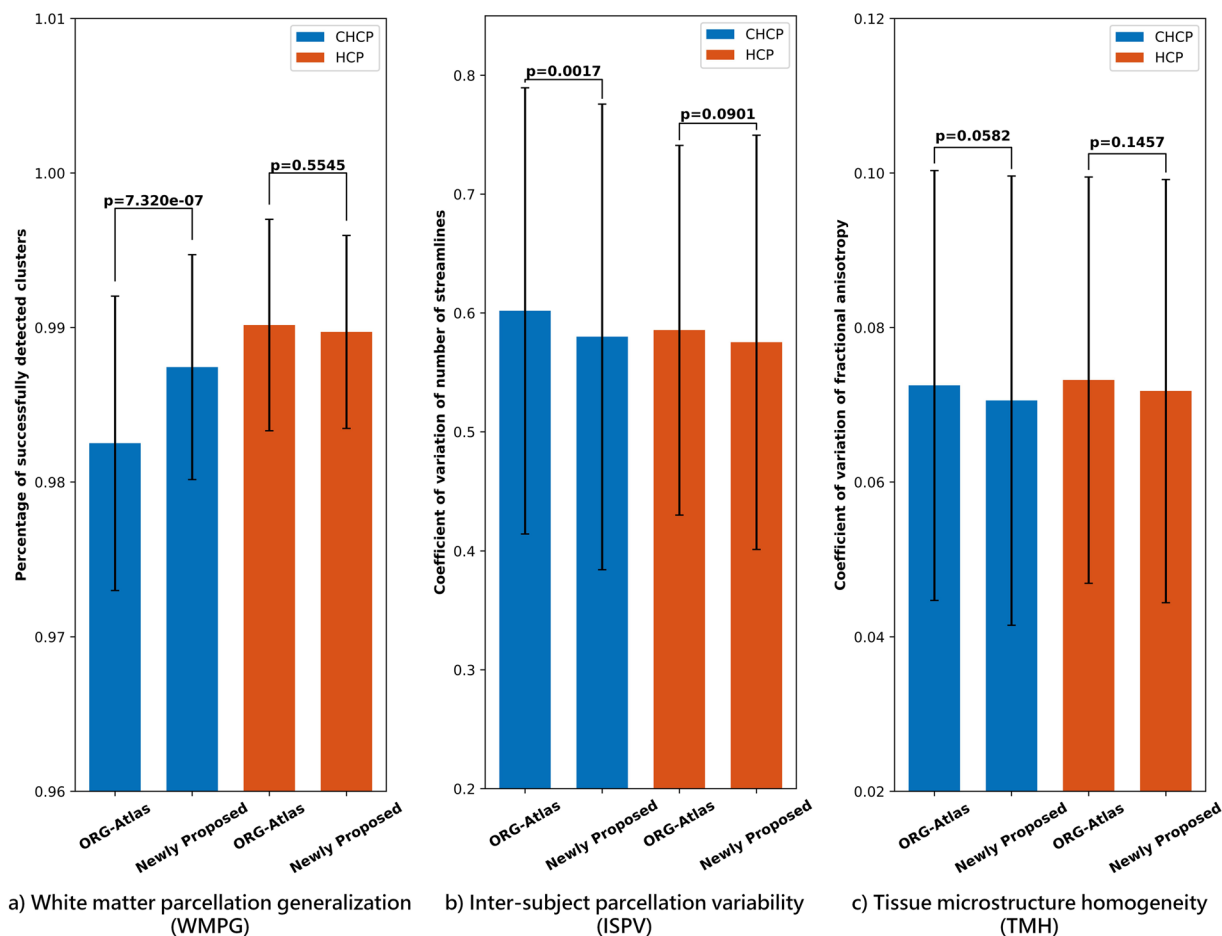
Fig. 8 Quantitative evaluation of whole brain WM parcellations in the CHCP and HCP-YA populations using the ORG atlas and our newly proposed atlas.

data alongside the HCP-YA dataset to enhance WM mapping in the CHCP population. On the other hand, it was worth noting that there were no significant differences in any of the three metrics in the HCP-YA data, underscoring the ability of the proposed atlas on concurrent WM mapping across the two populations.

Investigation of WM differences between the Eastern and Western populations. We first performed a group-wise comparison of the harmonized whole brain tractography data between the CHCP and HCP-YA data. The mean FA value of each subject's whole brain tractography data was computed, followed by an independent t-test for comparison between the CHCP and HCP-YA datasets. There was no significant difference between the two populations ($p = 0.70$). (It is worth noticing that there was a significant difference in the whole brain FA before harmonization, $p \ll 0.01$.) Then, we performed a group-wise comparison for each annotated anatomical fiber tract between the two populations under study. Specifically, for each fiber tract per subject, we extracted two measures of interest, i.e., connectivity strength that is measured as NoS of each tract (expressed as a proportion of the total NoS in the whole brain tractography) and the microstructure measure FA that reflects the diffusion anisotropy of the water molecules of the fiber tract. A general linear model (GLM) analysis was performed for each fiber tract, comparing the CHCP and HCP-YA groups, where the tract measure was the dependent variable, the group was a predictor variable, and age and gender were covariates. A Bonferroni method for multiple comparison corrections was performed across all the fiber tracts.

Figure 9 shows the results of the comparison between the two populations. In the connectivity strength comparison, there were 42 tracts with significant differences, in which 18 tracts had higher strengths in the CHCP group and 24 tracts had higher strengths in the HCP-YA group. In the FA comparison, there were 32 tracts with significant differences, in which 5 tracts had higher FA values in the CHCP group and 27 tracts had higher FA values in the HCP-YA group.

Usage Notes

In this study, we created a dMRI tractography atlas to enable concurrent mapping of WM connections in both Eastern and Western populations. Alongside the atlas, we also generated a large harmonized dMRI dataset from the two populations, including the diffusion imaging data, the whole brain tractography data, and parcellated WM fiber tracts and their diffusion measures. This resource serves as a valuable tool for exploring both
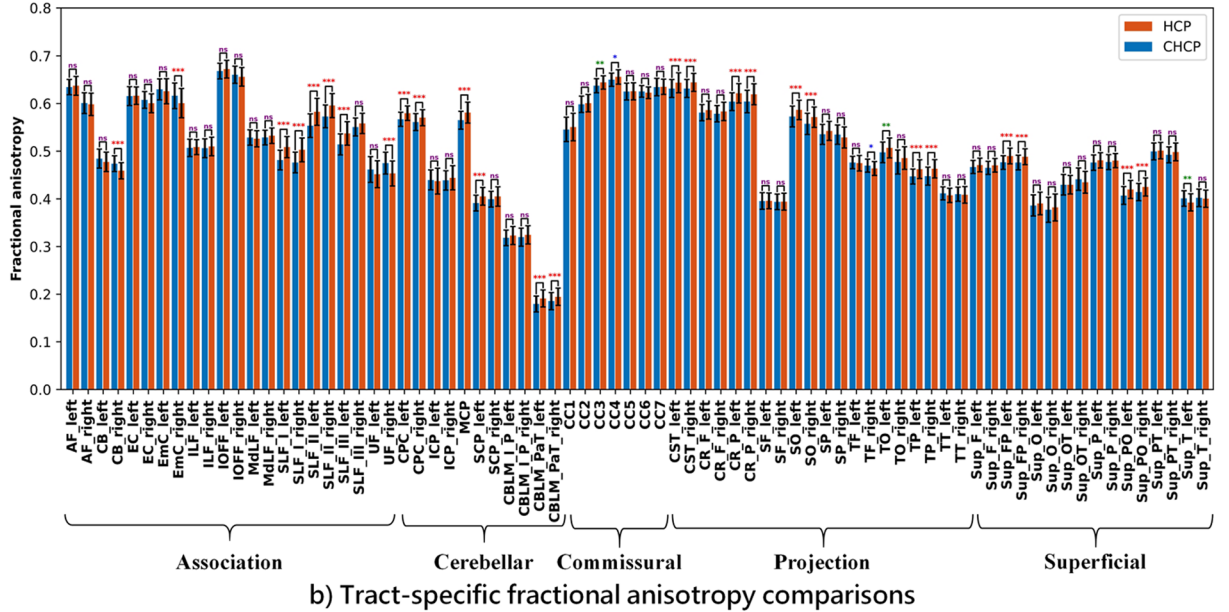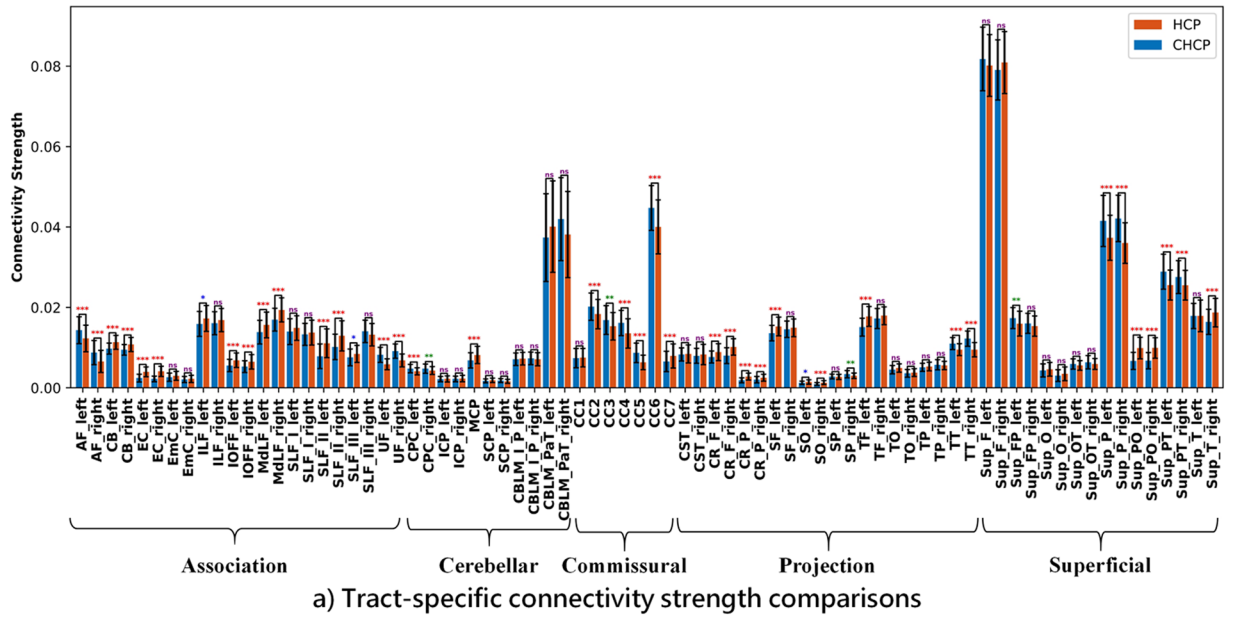
Fig. 9 (**a**) shows the comparison of the percentage of streamlines of each anatomical tract between CHCP and HCP-YA datasets. (**b**) shows the comparison of the FA values for each tract between CHCP and HCP-YA datasets. Notes: purple "*ns*" stands for no significance where p-value > 0.05, blue "*" stands for p-value < 0.05, green "**" stands for p-value < 0.01, and red "***" stands for p-value < 0.001.

commonalities and differences among diverse cultural backgrounds, offering openly accessible data to support further investigations. The potential usage of our work is discussed in the following aspects.

The subject-specific fiber tract data can be a valuable resource for conducting cross-cultural investigations into local WM regions associated with certain brain functions. These fiber tract data were derived from the harmonized dMRI dataset and the proposed cross-population atlas as a common template, enabling an unbiased comparison of corresponding WM structures between the two populations. Our initial exploration of the data shows significant differences between the two populations in several language-related fiber tracts such as AF, IOFF, UF, MdLF, SLF II, and ILF. This finding is consistent with the observed differences in language processing tasks between these cultural groups, as identified through functional MRI analyses[2,3,7,12,91,92]. Further investigation into fiber tracts related to other brain functions, such as motor and cognition, can elucidate potential differences in brain function, providing insights into the variations that may exist across cultural groups.

In addition to tract-specific analyses, the whole brain tractography data is available for conducting connectome-style analyses. The harmonized tractography data can serve as a valuable tool for further validating the results obtained in previous multi-site studies[2,21,22], eliminating the need for additional data harmonization processes. By providing a standardized and consistent dataset across populations, the harmonized tractography

enables a more direct comparison and evaluation of findings from diverse study sites. This contributes to the robustness and reliability of existing results, enhancing the overall validity of cross-cultural comparisons in brain research.

Moreover, the harmonized dMRI dataset empowers researchers to conduct a wide range of analyses to test their neuroscientific hypotheses concerning cross-cultural brain differences in dMRI. Unlike existing dMRI harmonization approaches that utilize meta-analysis or mega-analysis to aggregate data from different sites[93,94], our method harmonizes the dMRI data at the diffusion signal level. Beyond the aforementioned tractography-based analysis, the provided dMRI data can be seamlessly integrated into various downstream analysis methods. These include voxel-wise tract-based spatial statistics (TBSS)[95], voxel-based morphometry (VBM)[96] on standard diffusion tensor features, and advanced models like Neurite Orientation Dispersion and Density Imaging (NODDI)[97] and Free Water (FW)[98].

Lastly, the generated atlas and associated computational tools are publicly accessible, facilitating automated parcellation of WM tracts in new datasets. Through the harmonization of the new data to the atlas dMRI dataset, researchers can conduct WM parcellation based on the atlas. In this context, the proposed atlas serves as a valuable resource for cross-cultural comparisons in the context of brain diseases, aging, and neurodevelopment. The availability of this tool enhances the potential for comprehensive investigations into structural brain variations across diverse cultural contexts.

Because our study aims to create a tractography-based fiber tract atlas, we leverage the $b = 2000\,s/mm^2$ dMRI data with relatively high b values for improved fiber tracking. It is worth noting that this data is not ideal for diffusion models (e.g., diffusion tensor) that generally require lower b values. Therefore, diffusion metrics such as FA and MD derived from our dMRI data should not be directly compared to literature values reported from $b = 1000\,s/mm^2$ data. In this case result interpretation using our data should be performed considering the applications of interest. An interesting future direction is extending our data to include the $b = 1000\,s/mm^2$ to benefit any potential applications that require lower b values.

Code availability

The dMRI data processing software employed in this study is publicly available on GitHub. The repositories house the complete processing pipeline and scripts covering stages from preprocessing, CNN-based masking, dMRI harmonization, to whole-brain tractography, white matter parcellation, and visualization. These repositories offer comprehensive documentation, usage instructions, script examples on sample data, and list all dependencies necessary for effective utilization of the software. CHCP dMRI data processing pipeline: https://github.com/pnlbwh/pnlpipe/tree/v2.2.0; Convolutional neural network-based dMRI brain segmentation: https://github.com/pnlbwh/CNN-Diffusion-MRIBrain-Segmentation/tree/v0.3; dMRI data harmonization: https://github.com/pnlbwh/dMRIharmonization/tree/v2.1; UKF whole brain tractography: https://github.com/pnlbwh/ukftractography; White Matter Analysis: https://github.com/SlicerDMRI/whitematteranalysis; SlicerDMRI: http://dmri.slicer.org

References

1. Tooley, U. A., Bassett, D. S. & Mackey, A. P. Environmental influences on the pace of brain development. *Nature Reviews Neuroscience* **22**, 372–384 (2021).
2. Ge, J. *et al*. Increasing diversity in connectomics with the Chinese Human Connectome Project. *Nature Neuroscience* **26**, 163–172 (2023).
3. Kochunov, P. *et al*. Localized morphological brain differences between English-speaking Caucasians and Chinese-speaking Asians: new evidence of anatomical plasticity. *Neuroreport* **14**, 961–964 (2003).
4. Nisbett, R. E. & Miyamoto, Y. The influence of culture: holistic versus analytic perception. *Trends in cognitive sciences* **9**, 467–473 (2005).
5. Chua, H. F., Boland, J. E. & Nisbett, R. E. Cultural variation in eye movements during scene perception. *Proceedings of the national academy of sciences* **102**, 12629–12633 (2005).
6. Han, S. & Ma, Y. Cultural differences in human brain activity: a quantitative meta-analysis. *NeuroImage* **99**, 293–300 (2014).
7. Ge, J. *et al*. Cross-language differences in the brain network subserving intelligible speech. *Proceedings of the National Academy of Sciences* **112**, 2972–2977 (2015).
8. Gao, T., Han, X., Bang, D. & Han, S. Cultural differences in neurocognitive mechanisms underlying believing. *Neuroimage* **250**, 118954 (2022).
9. Tang, Y. *et al*. The construction of a Chinese MRI brain atlas: a morphometric comparison study between Chinese and Caucasian cohorts. *Neuroimage* **51**, 33–41 (2010).
10. Yang, G. *et al*. Sample sizes and population differences in brain template construction. *NeuroImage* **206**, 116318 (2020).
11. Kang, D. W. *et al*. Differences in cortical structure between cognitively normal East Asian and Caucasian older adults: a surface-based morphometry study. *Scientific reports* **10**, 20905 (2020).
12. Wei, X. *et al*. Native language differences in the structural connectome of the human brain. *Neuroimage* **270**, 119955 (2023).
13. Tang, Y. *et al*. Brain structure differences between C hinese and C aucasian cohorts: A comprehensive morphometry study. *Human brain mapping* **39**, 2147–2155 (2018).
14. Huang, C.-M., Doole, R., Wu, C. W., Huang, H.-W. & Chao, Y.-P. Culture-related and individual differences in regional brain volumes: A cross-cultural voxel-based morphometry study. *Frontiers in human neuroscience* **13**, 313 (2019).
15. Basser, P. J., Mattiello, J. & LeBihan, D. MR diffusion tensor spectroscopy and imaging. *Biophysical journal* **66**, 259–267 (1994).
16. Basser, P. J., Pajevic, S., Pierpaoli, C., Duda, J. & Aldroubi, A. In vivo fiber tractography using DT-MRI data. *Magnetic resonance in medicine* **44**, 625–632 (2000).
17. Zhang, F. *et al*. Quantitative mapping of the brain's structural connectivity using diffusion MRI tractography: A review. *Neuroimage* **249**, 118870 (2022).

18. Piper, R. J., Yoong, M. M., Kandasamy, J. & Chin, R. F. Application of diffusion tensor imaging and tractography of the optic radiation in anterior temporal lobe resection for epilepsy: a systematic review. *Clinical neurology and neurosurgery* **124**, 59–65 (2014).

19. Pannek, K., Scheck, S. M., Colditz, P. B., Boyd, R. N. & Rose, S. E. Magnetic resonance diffusion tractography of the preterm infant brain: a systematic review. *Developmental Medicine & Child Neurology* **56**, 113–124 (2014).

20. Essayed, W. I. *et al.* White matter tractography for neurosurgical planning: A topography-based review of the current state of the art. *NeuroImage: Clinical* **15**, 659–672 (2017).

21. Zhang, Y. *et al.* Structural connectivity profile supports laterality of the salience network. *Human Brain Mapping* **40**, 5242–5255 (2019).

22. Suo, X. *et al.* Anatomical and functional coupling between the dorsal and ventral attention networks. *Neuroimage* **232**, 117868 (2021).

23. Yeo, S. S., Jang, S. H. & Son, S. M. The different maturation of the corticospinal tract and corticoreticular pathway in normal brain development: diffusion tensor imaging study. *Frontiers in human neuroscience* **8**, 573 (2014).

24. Alexander, A. L. *et al.* Diffusion tensor imaging of the corpus callosum in Autism. *Neuroimage* **34**, 61–73 (2007).

25. Shany, E. *et al.* Diffusion tensor tractography of the cerebellar peduncles in prematurely born 7-year-old children. *The Cerebellum* **16**, 314–325 (2017).

26. Zekelman, L. R. *et al.* White matter association tracts underlying language and theory of mind: An investigation of 809 brains from the Human Connectome Project. *Neuroimage* **246**, 118739 (2022).

27. Ribeiro, M., Yordanova, Y. N., Noblet, V., Herbet, G. & Ricard, D. White matter tracts and executive functions: a review of causal and correlation evidence. *Brain* **147**, 352–371 (2024).

28. Laird, A. R. *et al.* ALE meta-analysis workflows via the brainmap database: progress towards a probabilistic functional brain atlas. *Frontiers in neuroinformatics* **3**, 598 (2009).

29. Mazziotta, J. *et al.* A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **356**, 1293–1322 (2001).

30. Collins, D. L., Neelin, P., Peters, T. M. & Evans, A. C. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *Journal of computer assisted tomography* **18**, 192–205 (1994).

31. Yang, G., Bozek, J., Han, M. & Gao, J.-H. Constructing and evaluating a cortical surface atlas and analyzing cortical sex differences in young Chinese adults. *Human brain mapping* **41**, 2495–2513 (2020).

32. Liang, P. *et al.* Construction of brain atlases based on a multi-center MRI dataset of 2020 Chinese adults. *Scientific reports* **5**, 18216 (2015).

33. Catani, M. & De Schotten, M. T. A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *cortex* **44**, 1105–1132 (2008).

34. Yendiki, A. *et al.* Automated probabilistic reconstruction of white-matter pathways in health and disease using an atlas of the underlying anatomy. *Frontiers in neuroinformatics* **5**, 23 (2011).

35. Román, C. *et al.* Clustering of whole-brain white matter short association bundles using HARDI data. *Frontiers in neuroinformatics* **11**, 73 (2017).

36. Zhang, F. *et al.* An anatomically curated fiber clustering white matter atlas for consistent white matter tract parcellation across the lifespan. *Neuroimage* **179**, 429–447 (2018).

37. Yeh, F.-C. *et al.* Population-averaged atlas of the macroscale human structural connectome and its network topology. *Neuroimage* **178**, 57–68 (2018).

38. Radwan, A. M. *et al.* An atlas of white matter anatomy, its variability, and reproducibility based on constrained spherical deconvolution of diffusion MRI. *NeuroImage* **254**, 119029 (2022).

39. Vollmar, C. *et al.* Identical, but not the same: intra-site and inter-site reproducibility of fractional anisotropy measures on two 3.0 T scanners. *Neuroimage* **51**, 1384–1394 (2010).

40. Mirzaalian, H. *et al.* Inter-site and inter-scanner diffusion MRI data harmonization. *NeuroImage* **135**, 311–323 (2016).

41. Karayumak, S. C. *et al.* Retrospective harmonization of multi-site diffusion MRI data acquired with different acquisition parameters. *Neuroimage* **184**, 180–200 (2019).

42. Cetin Karayumak, S., Kubicki, M. & Rathi, Y. Harmonizing diffusion MRI data across magnetic field strengths. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part III 11*, 116–124 (Springer, 2018).

43. Liu, Q., Dou, Q., Yu, L. & Heng, P. A. MS-Net: multi-site network for improving prostate segmentation with heterogeneous MRI data. *IEEE transactions on medical imaging* **39**, 2713–2724 (2020).

44. Ning, L. *et al.* Cross-scanner and cross-protocol multi-shell diffusion MRI data harmonization: Algorithms and results. *Neuroimage* **221**, 117128 (2020).

45. Tax, C. M. *et al.* Cross-scanner and cross-protocol diffusion MRI data harmonisation: A benchmark database and evaluation of algorithms. *NeuroImage* **195**, 285–299 (2019).

46. Mirzaalian, H. *et al.* Multi-site harmonization of diffusion MRI data in a registration framework. *Brain imaging and behavior* **12**, 284–295 (2018).

47. Pinto, M. S. *et al.* Harmonization of brain diffusion MRI: Concepts and methods. *Frontiers in Neuroscience* **14**, 485829 (2020).

48. Mirzaalian, H. *et al.* Harmonizing diffusion MRI data across multiple sites and scanners. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part I 18*, 12–19 (Springer, 2015).

49. Fortin, J.-P. *et al.* Harmonization of multi-site diffusion tensor imaging data. *Neuroimage* **161**, 149–170 (2017).

50. Huynh, K. M., Chen, G., Wu, Y., Shen, D. & Yap, P.-T. Multi-site harmonization of diffusion MRI data via method of moments. *IEEE transactions on medical imaging* **38**, 1599–1609 (2019).

51. Moyer, D., Ver Steeg, G., Tax, C. M. & Thompson, P. M. Scanner invariant representations for diffusion MRI harmonization. *Magnetic resonance in medicine* **84**, 2174–2189 (2020).

52. Hagler Jr, D. J. *et al.* Image processing and analysis methods for the Adolescent Brain Cognitive Development Study. *Neuroimage* **202**, 116091 (2019).

53. Van Essen, D. C. *et al.* The WU-Minn human connectome project: an overview. *Neuroimage* **80**, 62–79 (2013).

54. O'Donnell, L. J., Wells, W. M., Golby, A. J. & Westin, C.-F. Unbiased groupwise registration of white matter tractography. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012: 15th International Conference, Nice, France, October 1-5, 2012, Proceedings, Part III 15*, 123–130 (Springer, 2012).

55. O'Donnell, L. J. & Westin, C.-F. Automatic tractography segmentation using a high-dimensional white matter atlas. *IEEE transactions on medical imaging* **26**, 1562–1575 (2007).

56. Descoteaux, M., Angelino, E., Fitzgibbons, S. & Deriche, R. Regularized, fast, and robust analytical Q-ball imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* **58**, 497–510 (2007).

57. Ning, L. *et al.* Sparse Reconstruction Challenge for diffusion MRI: Validation on a physical phantom to determine which acquisition scheme and analysis method to use? *Medical image analysis* **26**, 316–331 (2015).

58. Glasser, M. F. *et al.* The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage* **80**, 105–124 (2013).

59. Cetin-Karayumak, S. *et al.* Harmonized diffusion MRI data and white matter measures from the Adolescent Brain Cognitive Development Study. *Scientific Data* **11**, 249 (2024).
60. Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W. & Smith, S. M. Fsl. *Neuroimage* **62**, 782–790 (2012).
61. Cetin-Karayumak, S. *et al.* Characterization of the extracellular free water signal in schizophrenia using multi-site diffusion MRI harmonization. *Molecular psychiatry* **28**, 2030–2038 (2023).
62. De Luca, A. *et al.* Cross-site harmonization of multi-shell diffusion MRI measures based on rotational invariant spherical harmonics (RISH). *NeuroImage* **259**, 119439 (2022).
63. Seitz-Holland, J. *et al.* Cognitive deficits, clinical variables, and white matter microstructure in schizophrenia: a multisite harmonization study. *Molecular psychiatry* **27**, 3719–3730 (2022).
64. Reisert, M., Kellner, E., Dhital, B., Hennig, J. & Kiselev, V. G. Disentangling micro from mesostructure by diffusion MRI: a Bayesian approach. *Neuroimage* **147**, 964–975 (2017).
65. Farquharson, S. *et al.* White matter fiber tractography: why we need to move beyond DTI. *Journal of neurosurgery* **118**, 1367–1377 (2013).
66. Vos, S. B., Viergever, M. A. & Leemans, A. Multi-fiber tractography visualizations for diffusion MRI data. *PloS one* **8**, e81453 (2013).
67. Zhang, F. *et al.* Deep white matter analysis (DeepWMA): Fast and consistent tractography segmentation. *Medical Image Analysis* **65**, 101761 (2020).
68. Xue, T. *et al.* Superficial white matter analysis: An efficient point-cloud-based deep learning framework with supervised contrastive learning for consistent tractography parcellation across populations and dMRI acquisitions. *Medical Image Analysis* **85**, 102759 (2023).
69. Zhang, F. *et al.* Creation of a novel trigeminal tractography atlas for automated trigeminal nerve identification. *Neuroimage* **220**, 117063 (2020).
70. Zeng, Q. *et al.* Automated identification of the retinogeniculate visual pathway using a high-dimensional tractography atlas. *IEEE Transactions on Cognitive and Developmental Systems* (2023).
71. O'Donnell, L. J. *et al.* Automated white matter fiber tract identification in patients with brain tumors. *NeuroImage: Clinical* **13**, 138–153 (2017).
72. Wu, Y., Ahmad, S. & Yap, P.-T. Highly reproducible whole brain parcellation in individuals via voxel annotation with fiber clusters. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 477–486 (Springer, 2021).
73. Zhang, F. *et al.* Test–retest reproducibility of white matter parcellation using diffusion MRI tractography fiber clustering. *Human brain mapping* **40**, 3041–3057 (2019).
74. Moberts, B., Vilanova, A. & Van Wijk, J. J. Evaluation of fiber clustering methods for diffusion tensor imaging. In *VIS 05. IEEE Visualization, 2005.*, 65–72 (IEEE, 2005).
75. Norton, I. *et al.* SlicerDMRI: open source diffusion MRI software for brain cancer research. *Cancer research* **77**, e101–e103 (2017).
76. Zhang, F. *et al.* SlicerDMRI: diffusion MRI and tractography research software for brain cancer surgery planning and visualization. *JCO clinical cancer informatics* **4**, 299–309 (2020).
77. Li, Y. *et al.* A diffusion MRI tractography atlas for concurrent white matter mapping across Eastern and Western populations. *ScienceDB* https://doi.org/10.57760/sciencedb.17354 (2024).
78. Gorgolewski, K. J. *et al.* The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific data* **3**, 1–9 (2016).
79. Fedorov, A. *et al.* 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic resonance imaging* **30**, 1323–1341 (2012).
80. Zhang, S. & Arfanakis, K. Evaluation of standardized and study-specific diffusion tensor imaging templates of the adult human brain: Template characteristics, spatial normalization accuracy, and detection of small inter-group FA differences. *Neuroimage* **172**, 40–50 (2018).
81. Qi, X. & Arfanakis, K. Regionconnect: Rapidly extracting standardized brain connectivity information in voxel-wise neuroimaging studies. *Neuroimage* **225**, 117462 (2021).
82. Fan, D. *et al.* Post-traumatic cerebral microhemorrhages and their effects upon white matter connectivity in the aging human brain. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 198–203 (IEEE, 2019).
83. Robles, D. J. *et al.* Older age, male sex, and cerebral microbleeds predict white matter loss after traumatic brain injury. *GeroScience* **44**, 83–102 (2022).
84. van Gool, R. *et al.* Characterization of central manifestations in patients with Niemann-Pick disease type C. *Genetics in Medicine* **26**, 101053 (2024).
85. Zanao, T. A. *et al.* Exploring the impact of hippocampal sclerosis on white matter tracts and memory in individuals with mesial temporal lobe epilepsy. *Epilepsia Open* **8**, 1111–1122 (2023).
86. Levitt, J. *et al.* The organization of frontostriatal brain wiring in non-affective early psychosis compared with healthy subjects using a novel diffusion imaging fiber cluster analysis. *Molecular Psychiatry* **28**, 2301–2311 (2023).
87. Cousineau, M. *et al.* A test-retest study on Parkinson's PPMI dataset yields statistically significant white matter fascicles. *NeuroImage: Clinical* **16**, 222–233 (2017).
88. Chen, Y. *et al.* Deep fiber clustering: Anatomically informed fiber clustering with self-supervised deep learning for fast and effective tractography parcellation. *NeuroImage* **273**, 120086 (2023).
89. Zhang, F. *et al.* Comparison between two white matter segmentation strategies: an investigation into white matter segmentation consistency. In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, 796–799 (IEEE, 2017).
90. Zhang, F. *et al.* DDParcel: deep learning anatomical brain parcellation from diffusion MRI. *IEEE Transactions on Medical Imaging* (2023).
91. Lu, J. *et al.* Functional maps of direct electrical stimulation-induced speech arrest and anomia: a multicentre retrospective study. *Brain* **144**, 2541–2553 (2021).
92. Wu, J. *et al.* Direct evidence from intraoperative electrocortical stimulation indicates shared and distinct speech production center between Chinese and English languages. *Human brain mapping* **36**, 4972–4985 (2015).
93. Kochunov, P. *et al.* Multi-site study of additive genetic effects on fractional anisotropy of cerebral white matter: comparing meta and megaanalytical approaches for data pooling. *Neuroimage* **95**, 136–150 (2014).
94. Zhang, G. *et al.* ENIGMA-Viewer: interactive visualization strategies for conveying effect sizes in meta-analysis. In *Proceedings of the 7th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, 640–648 (2016).
95. Smith, S. M. *et al.* Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage* **31**, 1487–1505 (2006).
96. Ashburner, J. & Friston, K. J. Voxel-based morphometry-the methods. *Neuroimage* **11**, 805–821 (2000).
97. Zhang, H., Schneider, T., Wheeler-Kingshott, C. A. & Alexander, D. C. NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage* **61**, 1000–1016 (2012).
98. Pasternak, O., Sochen, N., Gur, Y., Intrator, N. & Assaf, Y. Free water elimination and mapping from diffusion MRI. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* **62**, 717–730 (2009).

Acknowledgements

Author contributions

Conception and design: Yijie Li, Wei Zhang, Lauren J. O'Donnell, Fan Zhang; Data analysis and interpretation: Yijie Li, Wei Zhang, Suheyla Cetin-Karayumak, Kang Ik K Cho, Leo R. Zekelman, Lauren J. O'Donnell, Fan Zhang; Manuscript writing: All authors; Final approval of manuscript: All authors; Accountable for all aspects of the work: All authors.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to L.J.O. or F.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.