Individual Variability in Structural Connectivity Architecture of the Human Brain across the Middle-aged and Old Adult

Weijie Huang[1,2], Haojie Chen[2], Zhenzhao Liu[2], Guozheng Feng[2], Xinyi Dong[2], Li Su[3], GuoLing Ma[4], Zhanjun Zhang[2,5], Ni Shu[2,5]

[1]School of System Science, Beijing Normal University, Beijing, China

[2]State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China

[3]Department of Neuroscience, Neuroscience Institute, Insigneo Institute for in silico Medicine, University of Sheffield, Sheffield, United Kingdom

[4]Department of Radiology, China-Japan Friendship Hospital, Beijing, China

[5]BABRI Centre, Beijing Normal University, Beijing, China

Abstract

Human brain shows higher individual variability in functional connectivity in heteromodal association cortex but lower in unimodal cortex. Structural connectivity is the anatomical substrate of the functional connectivity. However, the spatial pattern of individual variability in structural connectivity still remains unclear. In the present study, we depicted the temporal-spatial pattern of individual variability in structural connectivity, which increases with aging and is lowest in unimodal sensorimotor regions, intermediate in association regions and highest in para-limbic regions. Our results demonstrated that this specific spatial distribution relates to the plasticity of synapse and white matter. Considering the brain systems and pathways involved in a cognition maybe different among individuals, we proposed a modified ridge regression model that is able to predict individual cognition and generate idiographic brain mapping. The individual variability in structural connectivity and idiographic

brain mapping is significantly correlated. Overall, our study has potential implications for understanding biological and genetic mechanisms of individual differences in structural connectivity, guiding interventions to promote successful cognitive and interpreting statistical maps in human connectome.

Introduction

Human thinks and behaves differently, which originates from interindividual variability in cerebral anatomy and function. A plethora of evidences suggest that such variability is heterogeneous across the cortex. For example, long association fibers, e.g., superior longitudinal fascicle, superior and inferior occipito-frontal fascicles are more variable than corticospinal tract, corpus callosum and optic tract in position and extent of fiber tracts [1]. In terms of cortical morphology, individual variability in cortical folding is higher in association areas than in motor cortex [2]. Functional network measuring the interregional neural synchronization act as an identifying fingerprint [3] and can be used to predict interindividual variation in behavior [4–7]. There is a growing interest in individual variability in functional connectivity (IVFC). Previous studies have revealed a stable pattern of such variability, with heteromodal association cortices' variability being significantly higher than that of unimodal cortices [8]. This pattern initially emerges in the third trimester resulting from heterogeneous changes of individual variability across different brain systems, with substantial decreases in sensorimotor system, intermediate decreases in visual, dorsal and ventral attention and subcortical systems, and slight decreases in default mode, frontoparietal and limbic systems [9]. In contrast to such decreases in development, IVFC increases with age in adults [10].

The potential causes of the spatial distribution of IVFC still remains unclear. A study found that IVFC is positively associated with evolutionary cortical expansion, which is partially retraced in individual development [8]. Therefore, the heterogeneity of IVFC across the cortex may be attributed to the different maturation trajectory. And those more variable areas such as association cortex is probably less genetically influenced because of the prolonged exposure to environment in maturation. However, A twin study found the complex but significant relationship between genes and IVFC that is age-dependent and bidirectional through systematic voxelwise analyses [11]. In addition, IVFC is positively associated with distant connectivity but negatively associated with

local connectivity has been consistently reported across studies [8–10].

The structural network derived from diffusion magnetic resonance imaging is the anatomical substrate of functional network. Nevertheless, the studies about individual variability in structural connectivity (IVSC) are scarce and their findings are also discordant. Two studies using the same public dataset, Human Connectome Project (HCP) showed a similar pattern of IVSC with inferior parietal, temporal, lateral prefrontal cortex and precuneus showing high variability [12,13]. But a study using a different dataset reported a different spatial distribution of IVSC that higher variability was localized near language areas, as well as in the medial frontal lobe, and in the superior frontal gyrus [14]. The heterogeneous findings may result from several factors: (a) the heterogeneity of population across studies; (b) methodological differences in the MRI acquisition parameters and methods to construct structural network. In addition, these studies all have some deficiencies in evaluating IVSC. For example, the size of the dataset in one study that only included nine participants is too small to obtain a reliable estimation of IVSC. Although another two studies using HCP had large dataset, they included twin subjects, which may lead to underestimation of IVSC. And none of these studies systematically explored the potential causes and genetical mechanisms of IVSC and the relationship between IVSC and cognition.

To address this gap, here we utilized multi-site datasets to characterize a reliable and replicable temporal-spatial pattern of IVSC. And we tested three related hypotheses. First, we hypothesized the IVSC is constrained by microscopic anatomical structure. Second, we expected that heterogeneity of IVSC across brain areas is modulated by genes. Third, we anticipated IVSC relates to individual differences in cognition.

Results

Temporal-spatial pattern of IVSC

To delineate the temporal spatial pattern of IVSC, 1734 participants from HCP Young Adults (HCP-YA) [15], HCP Retest [15], HCP Aging (HCP-A) [16], Cambridge Centre for Ageing and Neuroscience (Cam-CAN) [17] and Beijing Aging Brain Rejuvenation Initiative (BABRI) [18] were studies. Participants' demographical information was shown in **Table 1**.

Table 1. Demographical information of participants from 5 datasets.

Dataset	Sex (Male/Female)	Age (Year)	Trails A	Trails B	AVLT
HCP Retest	13/30	22.0 - 35.0 (30.3 ± 3.3)	-	-	-
HCP-YA	221/198	22.0 - 37.0 (28.3 ± 4.0)	-	-	-
HCP-A	202/274	36.0 - 67.9 (51.1 ± 9.1)	18.5 - 317.7 (66.5 ± 36.7)	0.3 -78.5 (26.9 ± 10.3)	10.0 - 74.0 (47.7 ± 10.0)
Cam-CAN	196/186	31.0 - 68.0 (49.4±11.1)	-	-	-
BABRI	142/272	45.0 - 68.0 (61.4 ± 8.6)	60.0 - 318.0 (146.0 ± 46.0)	24.0 - 115.0 (53.0 ± 14.1)	13.0 - 56.0 (31.0 ± 7.7)

HCP-YA, Human Connectome Project Young Adults; HCP-A, Human Connectome Project Aging; AVLT, auditory verbal learning task.

HCP Retest where participants underwent two MR scans within one year was used to estimate intraindividual difference in structural connectivity and regional regression coefficients between interindividual and intraindividual difference in structural connectivity. Small intraindividual difference in structural connectivity indicates high reliability of structural connectivity (**Figure S1A**). The intraindividual difference and regression coefficients were applied to the other datasets to estimate IVSC.

The IVSC of young adults (**Figure 1A**) from HCP-YA is similar to that of middle-aged and old adults from the combined dataset of HCP-A, Cam-CAN and BABRI (**Figure 1B**) (Spearman's $r = 0.808$, $p < 0.001$). Both IVSC maps show higher

variability in middle frontal cortex, inferior parietal cortex, middle and inferior temporal cortex, medial frontal cortex, medial temporal cortex and anterior cingulate cortex. To evaluate whether IVSC maps in middle-aged and old adults ware stable in different datasets, we harmonized structural networks from HCP-A, Cam-CAN and BABRI with ComBat [19] and calculated IVSC maps for the three datasets, respectively (**Figure S1B**, **S1C** and **S1D**). And the result shows that IVSC maps in middle-aged and old adults derived from different datasets are not only significantly associated to each other but also related to the IVSC map derived from HCP-YA and the combined dataset (**Figure 1C**) (All Spearman's r values > 0.75). Hence, the combined dataset was used for subsequent analyses to explore IVSC pattern across middle-aged and old participants. Additionally, we found that IVSC significantly increased with age in almost all brain by sliding window analysis (**Figure 1D**). Also, the mean IVSC and standard deviation of IVSC both positively correlates with age (**Figure 1E** and **1F**).

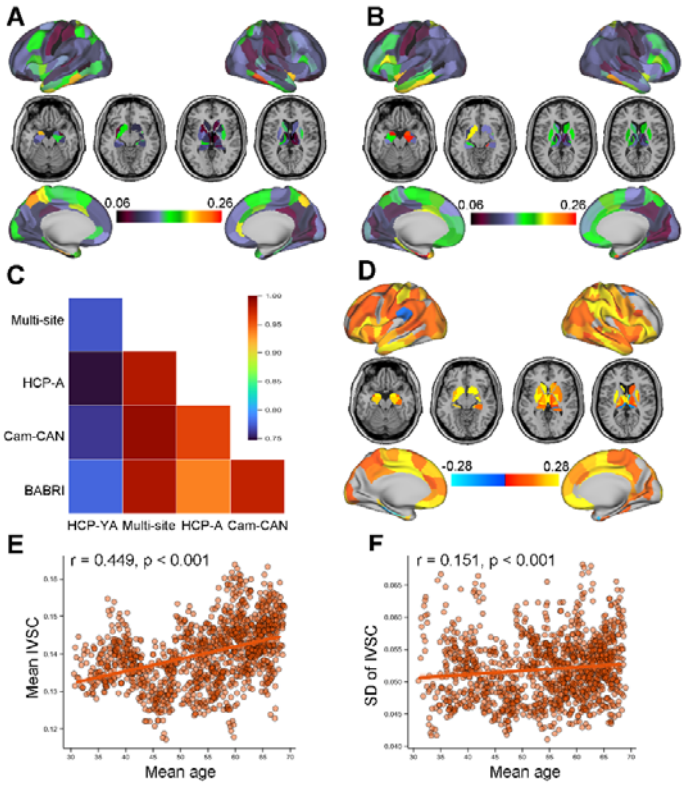

Figure 1. Temporal spatial pattern of IVSC. A, The IVSC across participants in HCP-YA. **B**, The IVSC across participants in HCP-A, Cam-CAN and BABRI. **C**, The correlation between IVSC from different datasets. **D**, Age-related rates of IVSC obtained by performing sliding

window analysis in the combined dataset of HCP-A, Cam-CAN and BABRI. **E**, The relationship between mean IVSC and mean age. **F**, The relationship between standard deviation of IVSC and mean age. IVSC, individual variability in structural connectivity; HCP-YA, Human Connectome Project Young Adults; HCP-A, Human Connectome Project Aging, Cam-CAN, Cambridge Centre for Ageing and Neuroscience; BABRI, Beijing Aging Brain Rejuvenation Initiative; SD, standard deviation.

IVSC was also assessed within 8 brain systems including 7 cortical networks [20] (**Figure 2A**) and a subcortical system. Limbic network had a high level of IVSC, whereas somatomotor network and visual network were least variable (**Figure 2B**). Sliding window analysis revealed that mean variability in all 8 brain systems significantly increased with age (**Figure 2C**).

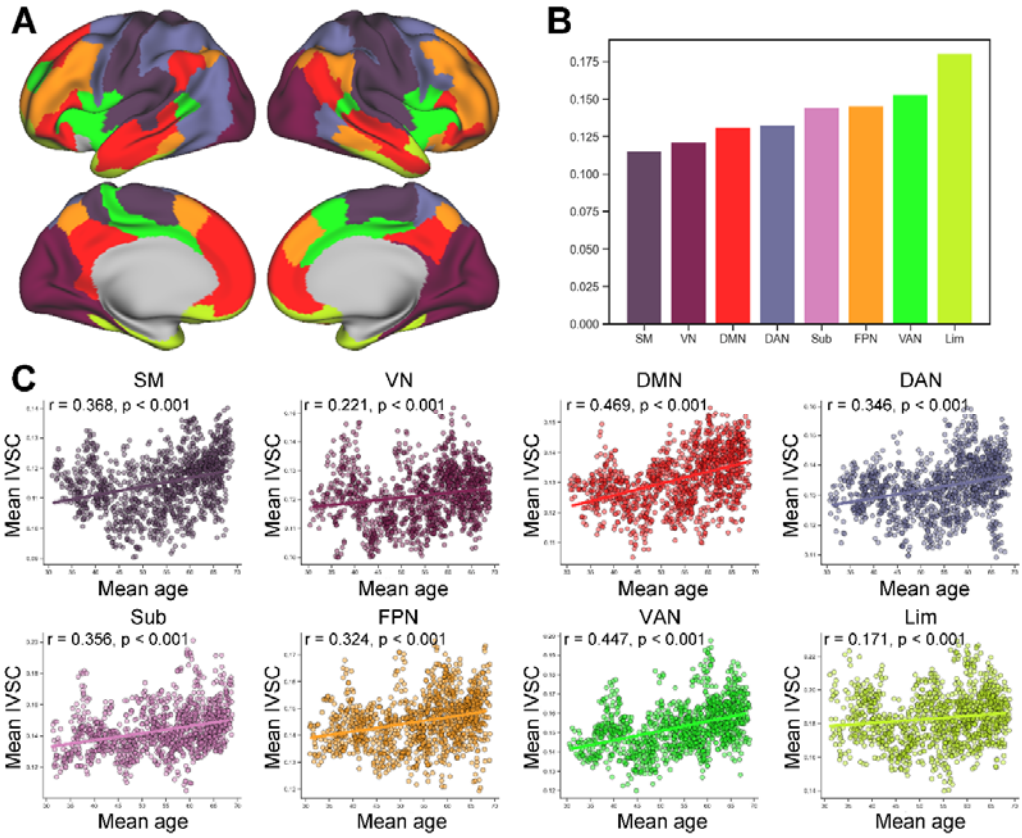

Figure 2. IVSC quantified in 8 brain systems. A, The spatial distribution of 7 cortical networks defined by Yeo and his colleagues [20]. **B**, Mean IVSC in 8 brain systems. **C**, The relationship between mean IVSC in 8 brain systems and mean age. IVSC, individual variability in structural

connectivity; SM, somatomotor network; VN, visual network; DMN, default mode network; DAN, dorsal attention network; Sub, subcortical system; FPN, frontosparietal network; VAN, ventral attention network; Lim, limbic network.

Association between IVSC and other properties of brain organization

Having identified the specific temporal spatial pattern of IVSC, we next sought to understand the biological mechanism underlying this specific spatial distribution of IVSC. The human cortex, no matter neocortex and allocortex, is consisted of several layers, which have a characteristic distribution of different neurons and connections with other cortical and subcortical. A previous study in mammals exhibited the existence of connection is not only associated with physical distance but also related to the cytology of cortical areas [21]. In addition, myelin proteins can inhibit axon sprouting and stop the formation of structural connection by collapsing the tips of growing axons [22]. Hence, we directly related IVSC to these fundamental properties of brain organization. Using spatial permutation testing, we found high IVSC was related to low laminar differentiation (**Figure 3D**; r = -0.461, p = 0.002), low myelin content (**Figure 3E**; r = -0.442, p < 0.001) and low strength of short SC (**Figure, 3F**; r = -0.421, p < 0.001).

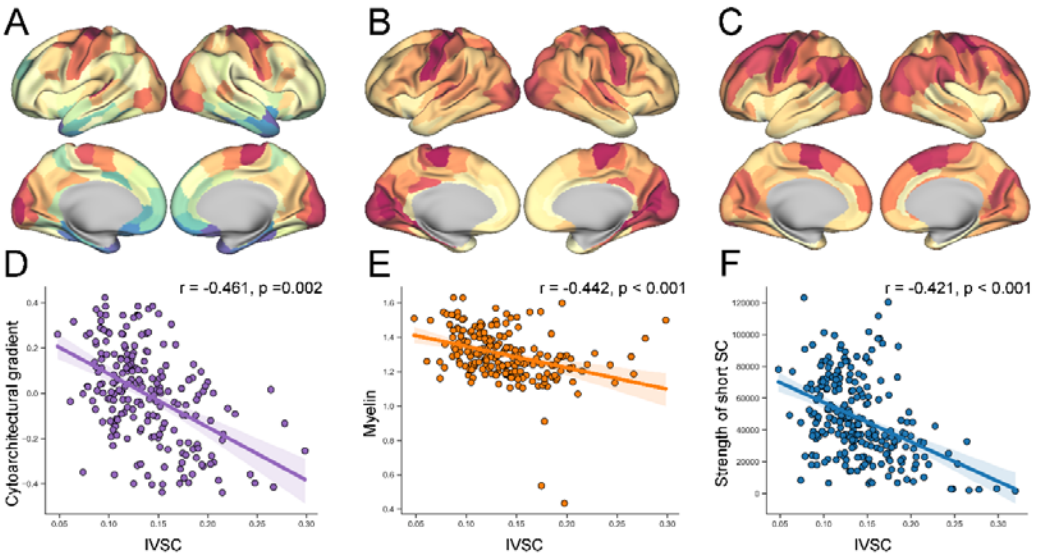

Figure 3. IVSC relates to fundamental properties of brain organization. A, Spatial distribution of cytoarchitectural gradient identified by Paquola et al [23]. **B**, Spatial distribution of

myelin content measured by T1w/T2w mapping [24]. **C**, Spatial distribution of short connectivity strength. **D-F**, IVSC was associated with cytoarchitecture gradient, myelin content and strength of short connectivity. IVSC, individual variability in structural connectivity; SC, structural connectivity.

Association between IVSC and gene expression

Next, we tried to explore the genetic mechanisms of IVSC using public transcriptional dataset. For Allen Human Brain Atlas (AHBA), after controlling the quality of gene expression and filtering nonbrain-specific genes, expression levels of 2112 brain-specific genes across 123 left brain regions were obtained. Using partial least square (PLS) analysis, a multi-variate statistical method to extract covarying components from matrices, we found a component of gene expression positively related to the IVSC **(Figure 4B**; r = 0.482, p = 0.005) and accounted for 26.1% of the IVSC's spatial variance. The component of gene expression is the sum of weighted gene expression, which we refer to as gene score. Medial temporal lobe, anterior cingulate cortex and insular showed highest gene scores while precentral gyrus, postcentral gyrus, and occipital showed lowest gene scores (**Figure 4A**).

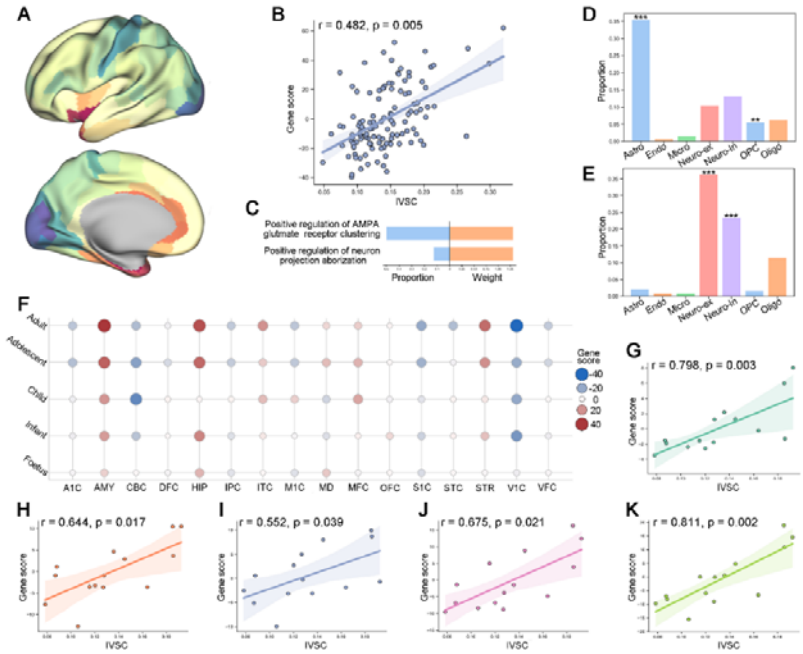

Figure 4. IVSC relates to gene expression profiles. A, Spatial distribution of gene expression

score identified by PLS. B, IVSC was correlated to gene expression score. C, Biological processes enriched in positive gene set were identified by performing gene set enrichment analysis and being tested against a spatial autocorrelation-preserving null model. D, The proportion of genes in each gene set preferentially expressed in seven different cell types. F, Gene scores through human development. G-K, IVSC was associated with gene score in foetus, infant, child, adolescent and adult. *$p < 0.05$, **$p<0.01$, ***$p<0.001$. IVSC, individual variability in structural connectivity; Astro, astrocyte; Micro, microglia; OPC, oligodendrocyte precursor; Oligo, oligodendrocyte; Endo, endothelial; Neuro-ex, excitatory neurons; Neuro-in, inhibitory neurons.

The relationship between IVSC and gene score is determined by the genes that contribute most to the gene score. The weights of genes reflect the shared variance between genes and gene score. We defined the genes with top 50% positive/negative weights as strongly contributing genes. To explore the biological process of the positive and negative gene sets, we performed gene set enrichment analyses on them, respectively. We identified two biological processes including “positive regulation of AMPA glutamate receptor clustering” ($p = 0.049$) and “positive regulation of neuron projection arborization” ($p = 0.049$) in positive gene set but no biological process in negative gene set (**Figure 4C**).

Having found the biological process enriched in positive gene set, we next investigated whether the genes associated with IVSC are preferentially expressed in specific cell types. Specifically, the proportion of genes that preferentially expressed in each cell type is calculated. Permutation tests were used to assess the statistical significance. Positive genes were significantly more expressed in astrocytes ($p < 0.001$) and oligodendrocyte precursor ($p = 0.004$) (**Figure 4D**) while negative genes were significantly more expressed in excitatory ($p < 0.001$) and inhibitory ($p < 0.001$) neurons (**Figure 4E**).

In addition, we used the BrainSpan dataset to evaluate the change of gene expression component identified by PLS through human development. The gene score of each

tissue sample was calculated based on the gene weights derived from above PLS analysis. The tissue samples were separated into five life stages (foetus, infant, child, adolescent and adult) in terms of their age. And the gene scores of tissue sample from the same life stage and same region were averaged to yield a gene score per region and life stage (**Figure 4F**). The gene scores in most regions increased with development. We also used the BrainSpan dataset to evaluate the reproducibility of association between IVSC and gene scores. We found IVSC significantly correlated to gene scores in the five life stages (**Figure 4G-K**; foetus: $r = 0.798$, $p = 0.003$; infant: $r = 0.644$, $p = 0.017$; child: $r = 0.522$, $p = 0.039$; adolescent: $r = 0.675$, $p = 0.021$, adult: $r = 0.811$, $p = 0.002$).

Individualized feature selection improves the prediction accuracy of structural connectivity

Considering that the variability in structural connectivity may arise from plastic change of white matter, the structural connectivity involved in cognitive tasks may be different across individuals. To verify the hypothesis, we modified the traditional ridge regression to the model that can select features individually to predict cognitions by integrating rectified linear unit activation function into ridge regression (see Methods for more details), and we investigated whether adding individualized feature selection into traditional machine learning method can improve accuracy of cognitive prediction by comparing the performance of traditional and modified ridge regression. Mean absolute error (MAE) and Pearson r value were used to evaluate the models' performance.

Structural connectivity can accurately predict executive function (**Figure 5A**; fold 1: MAE = 37.2, $p < 0.001$; Pearson $r = 0.54$, $p < 0.001$; fold 2: MAE = 34.5, $p < 0.001$; Pearson $r = 0.56$, $p < 0.001$) and attention (**Figure 5D**; fold 1: MAE = 10.2, $p < 0.001$; Pearson $r = 0.66$, $p < 0.001$; fold 2: MAE = 10.1, $p < 0.001$; Pearson $r = 0.66$, $p < 0.001$) in matched split-half samples with modified ridge regression. Repeated random 2-fold cross validation (2F-CV) returned similar results when predicting

executive function (**Figure 5B**, mean MAE = 37.9, mean Pearson r = 0.52) and attention (**Figure 5E**, mean MAE = 9.8, mean Pearson r = 0.67) and thus suggested that the matched split of the data was unbiased. We also compared the performance of modified and tradition ridge regression models derived from repeated random 2F-CV by using a two-sample t test and found then modified ridge regression models outperformed than traditional ridge regression models in both two tasks (**Figure 5B**; Executive function: T = -64.1, $p < 0.001$; **Figure 5E**; Attention function: T = -12.7, $p < 0.001$).

For each edge, we stat the number of participants who selected it as a feature when predicting executive function or attention. Then we defined the edge that were selected by at least two thirds of participants as consistent edge and the edge that were selected by less than a third of participants as individualized edge. Finally, we found the percentage of consistent edge was negatively associated with IVSC while the percentage of individualized edge was positively associated with IVSC when predicting executive function (**Figure 5C**; Consistent edge: Pearson r = -0.14, p = 0.025; Individualized edge: Pearson r = 0.18, p = 0.005) and attention (**Figure 5F**; Consistent edge: Pearson r = -0.23, $p < 0.001$; Individualized edge: Pearson r = 0.31, $p < 0.001$).

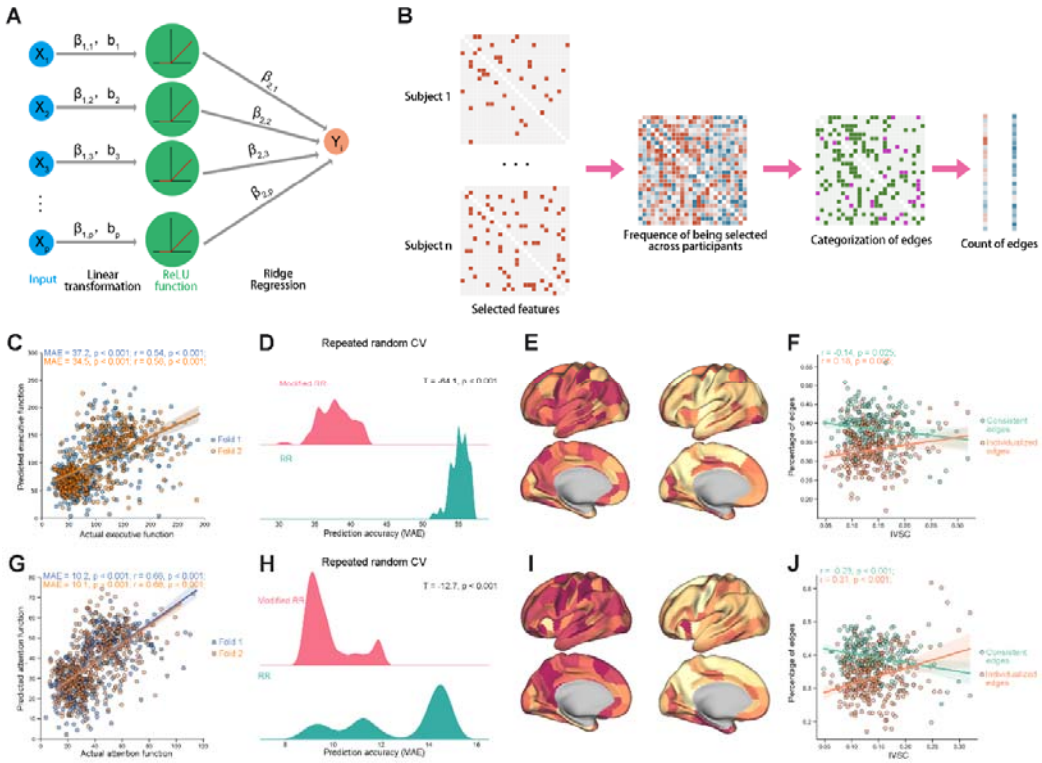

Figure 5. Prediction of executive function and attention benefits from individualized feature selection. A, Ridge regression was modified by integrating ReLU activation function into it. **B** For each edge, the number of participants who selected it as a feature was stated. Then the edges that were selected by at least two thirds of participants were defined as consistent edge sand the edges that were selected by less than a third of participants were defined as individualized edges. Finally, we calculate the percentage of consistent edges and individualized edges for each node. **C**, Structural connectivity can predict executive function with the modified ridge regression. **D**, Modified RR outperformed RR in the task of predicting executive function. **E**, Percentage of consistent edges (left) and individualized edges (right) in the prediction of executive function. **F**, Percentage of consistent edges and individualized edges related to IVSC across brain regions when predicting executive function. **G**, Structural connectivity can predict attention function with the modified ridge regression. **H**, Modified RR outperformed RR in the task of predicting attention function. **I**, Percentage of consistent edges (left) and individualized edges (right) in the prediction of attention. **J**, Percentage of consistent edges and individualized edges related to IVSC across brain regions when predicting attention function. ReLU, rectified linear unit. MAE, mean absolute error. CV, cross validation. RR, ridge regression. IVSC, individual variability in

structural connectivity.

Discussion

In this study, we leveraged a human connectome method and multi-center data to delineate the temporal-spatial pattern of IVSC and seek to explore the biological mechanism and significance of IVSC distribution. Building upon findings from this study, we identified a robust spatial distribution of IVSC with limbic network showing highest variability, association networks showing moderate variability and unimodal networks showing lowest variability. And IVSC increased in almost all brain with aging. In addition, we provided evidences that IVSC was strongly linked to microscopically anatomical properties of brain and gene expression. Finally, we demonstrated that individualized feature selection improved the prediction accuracy of structural connectivity. Taken together, these results offer a new perspective of brain organization and highlight the importance of individual effect in cognitive neuroscience research.

The spatial distribution of IVSC we observed was moderately different from previous studies although they all showed lowest variability in visual and somatomotor network [12–14]. The reason of inconsistent findings may mainly result from selection bias. The sample size of a previous study was only nine, which is too small to represent the population. In addition, decreased IVFC was associated with increased genetic sharing magnitude [11]. So, including twins in the other two studies may underestimate the IVSC especially in the regions with high heritability. Compared with FC, the interindividual variability in SC is lower on average than in FC in middle-aged and old adults and exhibits a different spatial pattern mainly in the limbic network (i.e. temporopolar and orbitofrontal area) which showing highest variability in SC but moderately lower variability in FC. Intuitively, this disassociation is possible because resting-state functional connectivity not only reflect the underlying anatomical circuitry but also recapitulates the history of experience-driven

coactivation [25,26].

Of note, the spatial distribution of IVSC is highly consistent with sensory-fugal gradient where one pole is occupied by the limbic areas, which are polarized towards the internal milieu and responsible for the regulation of memory, motivation, emotion and autonomic-endocrine function, and the opposite pole is primary sensory-motor areas, which area polarized towards extra-personal environment and provide a gateway for the entry of sensory information into the cerebral neural circuitry and for outputting motor program to bulbar and spinal motoneurons [27]. It was proposed based on anatomical, physiological and behavioral experiments in macaque monkeys and was demonstrated in human brain with cytoarchitectonic, functional imaging, electrophysiological recordings and the behavioral effects of focal lesions [27]. The areas further away from the external input have more variable white matter connections across human. In deep learning field, the connections in early and middle layers of architecture neural network is interpreted as a representation of the data and is considered generally applicable across similar tasks, but the connections in latter layers of architecture neural network is task-specific. Based on the theory, transfer learning, an approach to reuse of a pre-trained model on a new problem, is widely applied to a variety of scenarios. The shared structural characteristic between human brain network and architecture neural network may suggest the flexibility of cerebral latter layers, namely transmodal areas, is enough to guarantee the ability to adapt to changes of environment and the stability of sensor-motor areas play a critical role in the accurate registration of new inputs.

Having identified the spatial distribution of IVSC, we sought to understand whether the macroscopic pattern of structural connectivity is constrained by microscopic properties of brain structure. We observed that higher IVSC is associated with better laminar differentiation and less myelin content. Recent evidence has demonstrated that laminar differentiation related to the cortical plasticity-stability continuum of areas. The areas with worse laminar differentiation have more markers of synaptic

plasticity (i.e. glial fibrillary acidic protein) and less markers of stability (i.e. parvalbumin), while the areas with better laminar differentiation show the opposite trend [28]. In addition, intracortical myelin is another markers that limit synaptic plasticity, because the proteins in myelin including Nogo-A [29,30], MAG [31] and OMgp [32,33] can lead to the collapse of tips of growing axons and thus prevent the projections toward its target. The areas with higher plasticity are more easily influenced by life experience and may be more variable across populations. Hence, the heterogeneous spatial distribution of IVSC may be the macroscopic reflection of the difference in synaptic and axonal plasticity across different brain regions.

And the relationship between IVSC and the strength of short connectivity can suggest that the variation can be mainly sourced from the plastic change of weak short white matter connectivity. However, this sort of weak short connectivity may be noise because of the limitation of the tractography algorithm, although we used the most cutting-edge processing pipeline. We need the further study to prove it. We also found the IVSC increased with aging, which is consistent with the finding in the study of functional connectivity [10]. The age-related increase of IVSC indicates that the plastic change may last across lifespan. And the neuropathology of neurodegenerative disease may contribute part of increase of variation, although we only included relatively young and healthy participants.

To further understand the genetic mechanism underlying the distribution of IVSC, we used expression data from AHBA and found two enriched biological processes included “positive regulation of AMPA glutamate receptor clustering” and “positive regulation of neuron projection arborization”. AMPA receptor, a target of glutamate, has been proved its’ modulation is great of importance for the modifying synaptic efficacy [34]. Therefore, both biological processes are related to the change of neural pathways, which is anatomical infrastructure for supporting learning. In addition, we also observed that areas with higher IVSC are enriched for genetic signal of astrocytes and oligodendrocyte precursors. Interestingly, axonal electrical activity not only

mediates proliferation of oligodendrocyte precursors by controlling the production and release of the growth factors [35] but also triggers the release of cytokine leukemia inhibitory factor, which promotes myelination [36]. Combining these evidences, the areas with higher IVSC may have more expression of genes involving the formation of new pathway and the strengthening of existed pathway. Hence, the differences in synaptic and axonal plasticity across brain regions may result from the selective expression of the related genes.

Moreover, our results demonstrate that IVSC is linked to executive and attention function. We found that the regions with higher IVSC has more variable representations of cognition across individuals. This is against traditional neuroimaging study, where a statistical map was computed across a group of subjects. Specifically, traditional neuroimaging study is more likely to get a significant result in areas with low variability but less likely to get a significant result in areas with high variability. Hence, consideration of individual effect can help better understand individual differences in behavior.

Several methodological considerations in the present study should be noted. First, the structural networks we used to estimate IVSC were constructed based on a population-based atlas. But, the spatial topological properties of cortical regions vary across individuals [37–39]. So, using a population-based atlas may be adversed to estimate IVSC. Second, we regressed out intra-individual variability from inter-individual variability, but the intra-individual variability and inter-individual variability were estimated in different datasets. Third, the transcriptional data that we used are based on small samples of postmortem brains. More comprehensive microarray gene expression datasets are necessary for future studies.

In summary, we provide the evidence that IVSC is constrained by micro-anatomical property and gene expression. These findings may indicate that the pattern IVSC is the reflection of regional differences in plasticity and the result of selective expression

of related genes. Finally, our result also highlights the importance of individual effect when explain the individual difference in cognition.

Methods

Participants

The human MRI data employed in this work came from five datasets, HCP-YA [15], HCP Retest [15], HCP-A [16], Cam-CAN [17], BABRI [18]. Each study of the corresponding dataset was approved by the research ethics review boards of the respective institutions. This project was approved by the research ethics review board of the Beijing Normal University These five data sets are described below.

HCP-YA. We used the final release of HCP-YA that includes 1206 healthy young adult participants. We restricted our analysis to subjects for whom are from different families to avoid the effects of shared genes and environment on evaluation of individual variability (n = 457). After 38 subjects were excluded for a low quality in raw diffusion weighted MRI image or low quality in QSIPrep pipeline [40], the present study included a total of 419 healthy young adult participants from HCP-YA (221 males, age 22 - 37).

HCP Retest. Among the 1206 subjects from HCP-YA, 46 subjects underwent two separate scans with an interval ranging from 0.5 to 11 months. 3 subjects were excluded after quality control of MRI images and QSIPrep pipeline [40]. The two MRI scans of remaining 43 subjects from HCP Retest were included in this study (13 males, age 22 - 35).

HCP-A. HCP-A is an ongoing study aim to enroll 1200+ healthy adults (age 36 – 100+). We used the HCP-A Release 2.0 including 725 healthy adults. Considering the influence of gray matter atrophy due to aging on evaluation of IVSC, we excluded the participants who is over age of 68, which is a upper tertile of ages of all participants

from HCP-A, Cam-CAN and BABRI. After excluding the subjects whose raw MRI image or QSIPrep piplie [40] quality is low, 476 healthy adults from HCP-A were included in this study (202 males, age 36 - 68).

Cam-CAN. Cam-CAN enrolled 2681 healthy participants and 700 of those participants underwent MRI scans. Similarly, the participants over the age of 68 were excluded. We also excluded the participants under the age of 30 because we want to only focus on the IVSC in aging process. Finally, 382 participants from Cam-CAN were included in this study (196 males, age 31 - 68).

BABRI. BABRI is an ongoing study based on community cohort focusing on asymptomatic stages of dementia and aiming to develop prevention strategies for cognitive impairment. Similarly, after excluding participants over the age of 68, 414 healthy participants from BABRI were included (142 males, age 45 - 68).

Cognitive assessment

To explore the association between cognition with individual variability in aging process, we only included the cognitive data in HCP-A, Cam-CAN and BABRI. For the HCP-A dataset, in addition to NIH toolbox, additional cognitive tests including Rey auditory verbal learning task, delay discounting, trails A and B were administrated to all participants for cognitive assessments. For the Cam-CAN dataset, 14 behavioural tasks are employed to examine cognitive processing across five core domains: emotional processing, memory, motor and action function, executive function and language processing. Of the 14 tasks, hotel task was used to measure executive function and emotional memory and visual short-term memory provided the assessment of memory. For the BABRI dataset, a battery of neuropsychological tests was adopted to assess the cognitive function across five domains: memory, executive function, attention, visuospatial ability and language. Among the five domains, executive function was measured via Trails B, Stroop B and C test, attention was assessed by Trails A, symbol digit modalities and Stroop A test, and memory was

examined by Rey auditory verbal learning task.

Image acquisition

HCP-YA and HCP Retest. MRI data were collected with a customized 3T Connectome Scanner adapted from a Siemens Skyra. T1-weighted (T1w) scans used a 3D magnetization prepared rapid gradient echo (MPRAGE) (slices = 256, TR 2400 ms, TE = 2.14 ms, flip angle = 8°, voxel size = 0.7 mm isotropic). A multi-shell diffusion weighted echo-planar imaging sequence was used for diffusion-weighted MRI (dMRI) data (90 diffusion-weighted directions for each of b =1000, 2000 and 3000 s/mm^2 and 18 images with b = 0 s/mm^2, slices = 111, TR = 5500ms, TE = 89.50 ms, slice thickness = 1.25 mm, FOV = 210 × 180 mm^2, and acquisition matrix = 168 × 144)

HCP-A. MRI data were acquired with a 3T MRI scanner (Prisma, Siemens). T1w scans used a multi-echo MPRAGE (sagittal slices = 208, TR = 2500 ms, TE = 1.81/3.6/5.39/7.18 ms, TI = 1000 ms, slice thickness = 0.8 mm, flip angle = 8°, FOV = 256 × 256 mm^2, and acquisition matrix = 320 × 320). A multi-shell diffusion weighted echo-planar imaging sequence was used for dMRI data (92 diffusion-weighted directions for each of b =1500, 93 diffusion-weighted directions for 3000 s/mm^2, 14 images with b = 0 s/mm^2, slices = 92, TR = 3230ms, TE = 89.20 ms, slice thickness = 1.5 mm, FOV = 210 × 210 mm^2, and acquisition matrix = 140 × 140)

Cam-CAN. All MRI datasets were collected at a single site using a 3T MRI scanner (Trio, Siemens). A 3D MPRAGE was used for T1w scan (TR = 2250 ms, TE = 2.99ms, TI =900 ms, flip angle =9°, FOV = 256 mm × 240 mm × 192 mm, voxel size = 1 mm isotropic). A twice-refocused spin-echo sequence was used for dMRI (30

diffusion gradient directions for each of b = 1000 and 2000 s/mm^2, 3 images with b = 0 s/mm^2, axial slices = 66, TR = 9100 ms, TE = 104 ms, FOV = 192 mm × 192 mm)

BABRI. MRI data were acquired with a 3T MRI scanner (Trio, Siemens Magnetom). A sagittal 3D MPRAGE was used for T1w scan (sagittal slices = 176, TR = 1900 ms, TE = 3.44 ms, TI = 900 ms, slice thickness = 1 mm, flip angle = 9°, FOV = 256 × 256 mm^2, and acquisition matrix = 256 × 256). A single-shell diffusion weighted echo-planar imaging sequence was used for dMRI data (30 diffusion-weighted directions with b = 1000 s/mm^2 and a single image with b = 0 s/mm^2, axial slices = 70, TR = 9500 ms, TE = 92 ms, slice thickness = 2 mm, FOV = 256 × 256 mm^2, and acquisition matrix = 128 × 128).

Image processing

The image processing pipeline consisted of preprocessing and reconstruction workflows. QSIPrep [40] 0.14.3 that is based on FSL [41], DSI studio [42], DIPY [43], ANTs [44]and MRtrix [45] is used to perform image processing and create individual white matter network. Detailed procedures are described as follows.

Preprocessing. For T1w images, ANTs was used to extract brain and generate brain mask. FSL was utilized for brain tissue segmentation. Finally, ANTs was used to normalize T1w image to MNI space. For dMRI images, images were grouped according to there phase encoding direction firstly. Then, MRtrix was used to denoise diffusion-weighted images. FSL was used for distortion and head motion correction. Finally, the template image with b = 0 was created and registered to the skull-stripped T1w image.

Reconstruction. Reconstruction workflow performed advanced reconstruction and tractography methods to create white matter network. Single-shell and multi-shell

constrained spherical deconvolution were separately applied to single-shell and multi-shell dMRI data to estimate voxel-level diffusion models. A probabilistic tractography algorithm in MRtrix was used to produces a fixed number of streamlines and anatomically-constrained tractography framework was used during tracking. Then, spherical-deconvolution informed filtering of tractograms was applied to improve the quantitative nature of whole-brain steamlines reconstructions. Finally, Brainnetome atlas [46] was used to define nodes in white matter network and number of streamlines connecting a pair of regions was used to weight edges.

Harmonization of multi-sites data

For removing the site effects in structural connectivity from multiple sites, ComBat model that was successfully applied to the harmonization of the measures from MRI data, such as cortical thickness [19] and function connectivity [47] was utilized. It assumes that sites have both an additive and multiplicative effects on data. Hence, the model is defined as:

$$y_{ijv} = \alpha_v + X_{ij}^T \beta_v + \gamma_{iv} + \delta_{iv} \epsilon_{ijv},$$

where α_v is the mean value of structural connectivity v; X_{ij}^T is a design matrix for the covariates β_v; γ_{iv} and $\delta_{iv}$are the additive and multiplicative site effects of site i for connectivity v, respectively. The final ComBat-harmonized connectivity are then defined as:

$$y_{ijv}^{ComBat} = \frac{y_{ijv} - \widehat{\alpha_v} - X_{ij}\widehat{\beta_v} - \gamma_{iv}^*}{\delta_{iv}^*} + \widehat{\alpha_v} + X_{ij}\widehat{\beta_v},$$

where γ_{iv}^* and δ_{iv}^* are the estimates of γ_{iv} and δ_{iv} using empirical Bayes. In this study, we included age and sex as covariates to preserve biological information in the data.

Definition of individual variability of structural network

The individual variability of structural network was quantified according to the definition in previous studies [8–10]. For a given node i, the inter-subject difference was

defined as follows:

$$R_i = 1 - E\left[corr\left(C_i(s_j, t_1), C_i(s_k, t_1)\right)\right],$$

where j, k = 1,2 … M (j ≠ k); M is the number of participants; $C_i(s_j, t_1)$ is a vector of structural connectivity values of node i in participant j at timepoint t_1.

The intra-subject difference was estimated using HCP Retest dataset where each participant has two MRI scans. It was defined as follows:

$$N_i = 1 - E\left[corr\left(C_i(s_j, t_1), C_i(s_j, t_2)\right)\right].$$

Then, individual variability of structural network is defined as follows:

$$V_i = R_i - \beta N_i - c,$$

where β and c are parameters determined via ordinary least-squares in HCP Retest dataset. They and N were applied to other datasets to estimate individual variability of structural network.

Sliding window analysis

To examine the aging of structural network variability, cross-subject sliding window analyses were performed for participants from HCP-A, Cam-CAN and BABRI in ascending order of age, with window length of 10 participants and a step size of 1 participant. Using this method, 1253 overlapping subgroups were generated and their structural network variabilities and mean ages were calculated.

Spatial permutation testing (Moran spectral randomization)

To evaluate the association between the spatial distribution of structural network variability and other cerebral properties. A spatial permutation procedure called Moran spectral randomization was used to control the effect of spatial auto-correlation on the association. Inverse Euclidean distance between nodes was used to denote the spatial proximity when using Moran spectral randomization generated 10000 random brain maps that preserved spatial auto-correlation of the

original map. The distribution of Spearman R values was obtained by relating structural network variability with random brain maps. The P value is computed as the proportion of null R value that are greater than true R value.

Transcriptional data preprocess

Transcriptional data of six postmortem brains were provided by the Allen Human Brain Atlas [48]. Analyses were limited to left brain only, because of only two of six brains including samples from right hemisphere. Abagen toolbox [49] was used to preprocess the transcriptional data according to the suggestions in a previous study [50].

Specifically, re-annotate microarray probes to corresponding genes based on the database provided by Arnatkevičiūtė et al. [51]. Filter noisy probes that do not exceed the background in at least 50% all samples across all subjects. For each gene, select the representative probe with expression level most correlated to RNA-seq measures collect in two 2 of six brains. Next, tissue samples were assigned to the closest regions in Brainnetome atlas, while the samples more than 2 mm from all regions would not be assigned to any regions. To account for inter-individual variation, expression values within each brain across regions were normalized with a scaled robust sigmoid function [52]. To retain the genes with consistent expression pattern across six brains, differential stability [53] was calculated for each gene and a threshold of 0.1 was imposed. Finally, regional expression values were averaged across donors.

It's important to provide gene specificity on observed transcriptomic-neuroimaging effect for avoiding nonspecific effects [54]. To retain brain-specific genes, genes with higher expressions in brain sites than non-brain sites were selected using Genotype-Tissue Expression database (https://www.gtexportal.org/). This procedure produced a matrix with 246 rows corresponding to brain regions and 2112 columns corresponding to the retained genes.

PLS analysis

The correlation between gene expression and structural network variability was examined with PLS analysis that is an unsupervised multivariate method to project two matrices into a lower-dimensional space where the covariance between of them is maximal. This analysis was performed using the *PLSRegression* function in Python package scikit-learn (https://scikit-learn.org/stable/index.html) with setting number of components as 1.

Bootstrap resampling

To obtain reliable estimation of gene weights, bootstrap resampling was used. The values in structural network variability vector were randomly selected with replacement 10000 times. PLS analysis was performed using these bootstrapped structural network variability vectors resulting in distributions of gene weights. The reliable gene weights were quantified by the ratio of the true gene weight to the standard error of its bootstrapped distribution. The ratio is the combination of both value and reliability of gene weights.

Gene set enrichment analysis

To further investigate biological pathways that related to the genes identified by PLS analysis, the method used by Hansen et al. [55] was adapted. The latest gene annotation) provided by Gene Ontology (geneontology.org) was downloaded on 2nd July 2022. For each biological process in gene annotation, the enrichment score was defined as the mean weights of the genes in the biological process. The null distribution of enrichment scores was constructed by permuting structural network variability map 10000 times with Moran spectral randomization and repeated PLS analysis and the calculation of enrichment score. This analysis was separately performed on the positive genes with the 50% most positive weights and negative genes with the 50% most negative weights. For positive genes, significance was measured by the proportion of null enrichment scores that are greater than true enrichment score. For negative genes, the proportion of null enrichment scores that are smaller than true enrichment score was utilized to quantify significance.

To determine cell types where the genes identified by PLS preferentially express, the method performed by Hansen et al. [55] was adopted. Briefly, gene annotation about cell type was provided by Seidlitz et al. [56], who used hierarchical clustering based on spatial expression profiles to distinguish 7 canonical cell types: astrocytes, endothelial cells, microglia, excitatory neurons, inhibitory neurons, oligodendrocytes and oligodendrocyte precursors. For each cell type, enrichment score was defined as the proportion of genes that were identified by PLS analysis and preferentially express in that type as well. This analysis was separately performed on the positive genes with the 50% most positive weights and negative genes with the 50% most negative weights. And non-parameter permutation test was used to quantify the significance.

Development of gene expression

We used BrainSpan [57], a database of gene expression across development, to depict the development of the gene expression component determined by PLS analysis in Allen Human Brain Atlas. Postmortem brains in BrainSpan were grouped into 5 developmental stages: foetus (8 – 37 post-conception weeks), infant (4 months – 1 year), child (2 – 8 year), adolescent (11 – 19 years) and adult (21 – 40 years). For each group, gene expression scores in 16 regions were calculated by multiplying the gene weights derived from the PLS analysis in Allen Human Brain Atlas and mean expression across postmortem brains in the group.

To relate gene expression scores in BrainSpan and structural network variability, we mapped Brainnetome Atlas into 16 Brodmann areas that are included in BrainSpan. For each Brodmann area, structural network variability was averaged across the Brainnetome-Atlas regions corresponding to the Brodmann area. Then the relationship between gene expression scores and structural network variability was examined in each developmental stage.

Modified ridge regression

To test whether the structural connectivity involved in cognitive tasks are different across individuals, we used two models to predict individual cognition. One is ridge regression formulized as follows:

$$y_i = \sum_{j=1}^{p} \beta_j x_{i,j} + \beta_0 + \varepsilon_i,$$

where y_i is the cognitive score of individual i, $x_{i,j}$ is the weight of structural connectivity j of individual i, β_j is the regression coefficient and ε_i is the error. Ridge regression has been showed to have similar performance to other machine learning methods but computationally more efficient when using high-dimensional imaging data as features [5]. When fitting the model, ridge regression used an L2 penalty to avoid over-fitting and to improve generalizability. The objective function is:

$$min_\beta \sum_{i=1}^{N} (f(x_i) - y_i)^2 + \alpha \sum_{j=1}^{p} \|\beta_j\|^2,$$

where α is a hyper-parameter that controls the penalty of model complexity.

Considering that the structural connectivity encoding cognitions may be different across individuals, we proposed a modified ridge regression that is a neural network with a hidden layer. The model can was defined as follows:

$$y_i = \sum_{j=1}^{p} \beta_{2,j} * ReLU(\beta_{1,j} * x_{i,j} + \beta_{1,0}) + \beta_{2,0} + \varepsilon_i,$$

where $\beta_{1,j}$ is the coefficient to linearly transform structural connectivity, $\beta_{2,j}$ is the coefficient to predict cognition and $ReLU(\cdot)$ is a linear rectification function. The $ReLU(\cdot)$ and $\beta_{1,j}$ together determine which features are used to predict cognition. Similarly, the modified ridge regression's objective function is:

$$min_\beta \sum_{i=1}^{N} (f(x_i) - y_i)^2 + \alpha(\sum_{j=1}^{p} \|\beta_{1,j}\|^2 + \sum_{j=1}^{p} \|\beta_{2,j}\|^2).$$

For each connectivity in each participant, the degree of that the connectivity was involved in a cognition can be defined as follows:

$$ReLU(x_{i,j} * \beta_{1,j} + \beta_{1,0}) * w_{2,j}.$$

Prediction framework

To train models of cognitive prediction, a nested 2F-CV where outer 2F-CV was for evaluation of models' performance and inner 2F-CV was used to determine the optimal hyper-parameter α.

Outer 2F-CV. In the outer 2F-CV, we merged HCP-A and BABRI datasets, because they both used the Trails B, Trails A and Rey auditory verbal learning to measure individuals' executive function, attention and memory respectively. Then participants in the merged dataset were sorted according to the outcome (i.e., executive function or attention). The participants with an odd rank were assigned to subset 1 and other participants were assigned to subset 2. In the first fold, subset 1 was used as training dataset and subset 2 was used as test dataset. The second fold is opposite. Each feature was normalized across the training dataset, and the mean value and standard deviation derived from training dataset were also applied to test dataset. After optimal hyper-parameter α was selected, a model with the optimal α was trained using all participants in the training dataset and then was applied to test dataset to predict the outcome. Pearson correlation coefficients and mean absolute error between predicted and true outcome of participants in test dataset was calculated to quantify prediction accuracy.

Inner 2F-CV. In the inner 2F-CV, the training dataset was further split into 2 subsets similarly to the outer 2F-CV. They were used to train models under different values of hyper-parameter α ([2-4, 2-3, …, 24, 25]) in turn, and these models were applied to remaining dataset to calculate Pearson correlation coefficient and mean absolute error between predicted and true outcome. For each value of α, Pearson correlation coefficients and mean absolute errors was averaged across 2 inner loops. Then the sum of normalized mean Pearson correlation coefficient and normalized reciprocal of the mean MAE was defined as inner prediction accuracy. The value of

hyper-parameter α with the highest inner prediction accuracy was selected as the optimal α.

Randomly split 2F-CV. To validate whether the result is influenced by data split, we randomly split the merged dataset into two halves in both outer 2F-CV and inner 2F-CV, and calculated the mean Pearson correlation coefficient and mean absolute error across two outer folds. This procedure was repeated 100 times to get the distribution of mean Pearson correlation coefficients and mean absolute errors.

Significance of prediction accuracy. A permutation test was performed to evaluate significance of prediction accuracy. Briefly, we performed the nested 2F-CV 1000 times with permuting outcome across training dataset in each run. Null distributions of Pearson correlation coefficients and mean absolute errors were obtained. The significance of Pearson correlation coefficient was determined as the proportion of null coefficient higher than true coefficient and significance of mean absolute error was defined as the proportion of null values lower than true value.

Reference

1. Bürgel, U. *et al.* White matter fiber tracts of the human brain: Three-dimensional mapping at microscopic resolution, topography and intersubject variability. *NeuroImage* **29**, 1092–1105 (2006).
2. Hill, J. *et al.* A Surface-Based Analysis of Hemispheric Asymmetries and Folding of Cerebral Cortex in Term-Born Human Infants. *Journal of Neuroscience* **30**, 2268–2276 (2010).
3. Finn, E. S. *et al.* Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat Neurosci* **18**, 1664–1671 (2015).
4. Rosenberg, M. D. *et al.* A neuromarker of sustained attention from whole-brain functional connectivity. *Nat Neurosci* **19**, 165–171 (2016).
5. Cui, Z. & Gong, G. The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features. *NeuroImage* **178**, 622–637 (2018).
6. Beaty, R. E. *et al.* Robust prediction of individual creative ability from brain functional connectivity. *Proceedings of the National Academy of Sciences* **115**, 1087–1092 (2018).
7. Shen, X. *et al.* Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *Nat Protoc* **12**, 506–518 (2017).
8. Mueller, S. *et al.* Individual Variability in Functional Connectivity Architecture of the Human Brain. *Neuron* **77**, 586–595 (2013).
9. Xu, Y. *et al.* Development and Emergence of Individual Variability in the

Functional Connectivity Architecture of the Preterm Human Brain. *Cerebral Cortex* **29**, 4208–4222 (2019).

10. Ma, L., Tian, L., Hu, T., Jiang, T. & Zuo, N. Development of Individual Variability in Brain Functional Connectivity and Capability across the Adult Lifespan. *Cerebral Cortex* bhab059 (2021) doi:10.1093/cercor/bhab059.
11. Gao, W. *et al.* Intersubject Variability of and Genetic Effects on the Brain's Functional Connectivity during Infancy. *Journal of Neuroscience* **34**, 11288–11296 (2014).
12. Sun, L. *et al*. Structural insight into the individual variability architecture of the functional brain connectome. 2022.02.16.480803 Preprint at https://doi.org/10.1101/2022.02.16.480803 (2022).
13. Mansour L, S., Tian, Y., Yeo, B. T. T., Cropley, V. & Zalesky, A. High-resolution connectomic fingerprints: Mapping neural identity and behavior. *NeuroImage* **229**, 117695 (2021).
14. Chamberland, M. *et al.* On the Origin of Individual Functional Connectivity Variability: The Role of White Matter Architecture. *Brain Connect* **7**, 491–503 (2017).
15. Van Essen, D. C. *et al*. The WU-Minn Human Connectome Project: an overview. *Neuroimage* **80**, 62–79 (2013).
16. Harms, M. P. *et al*. Extending the Human Connectome Project across ages: Imaging protocols for the Lifespan Development and Aging projects. *NeuroImage* **183**, 972–984 (2018).

17. Cam-CAN *et al*. The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) study protocol: a cross-sectional, lifespan, multidisciplinary examination of healthy cognitive ageing. *BMC Neurol* **14**, 204 (2014).

18. Yang, C. *et al*. Early prevention of cognitive impairment in the community population: The Beijing Aging Brain Rejuvenation Initiative. *Alzheimer's & Dementia* alz.12326 (2021) doi:10.1002/alz.12326.

19. Fortin, J.-P. *et al.* Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage* **167**, 104–120 (2018).

20. Thomas Yeo, B. T. *et al*. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* **106**, 1125–1165 (2011).

21. Goulas, A., Majka, P., Rosa, M. G. P. & Hilgetag, C. C. A blueprint of mammalian cortical connectomes. *PLoS Biol* **17**, e2005346 (2019).

22. Fields, R. D. White matter in learning, cognition and psychiatric disorders. *Trends in Neurosciences* **31**, 361–370 (2008).

23. Paquola, C. *et al*. Microstructural and functional gradients are increasingly dissociated in transmodal cortices. *PLoS Biol* **17**, e3000284 (2019).

24. Glasser, M. F. & Essen, D. C. V. Mapping Human Cortical Areas In Vivo Based on Myelin Content as Revealed by T1- and T2-Weighted MRI. *J. Neurosci*. **31**, 11597–11616 (2011).

25. Lewis, C. M., Baldassarre, A., Committeri, G., Romani, G. L. & Corbetta, M. Learning sculpts the spontaneous activity of the resting human brain. *Proceedings of the National Academy of Sciences* **106**, 17558–17563 (2009).

26. Tambini, A., Ketz, N. & Davachi, L. Enhanced Brain Correlations during Rest Are Related to Memory for Recent Experiences. *Neuron* **65**, 280–290 (2010).

27. Mesulam, M. From sensation to cognition. *Brain* **121**, 1013–1052 (1998).

28. García-Cabezas, M. Á., Joyce, M. K. P., John, Y. J., Zikopoulos, B. & Barbas, H. Mirror trends of plasticity and stability indicators in primate prefrontal cortex. *European Journal of Neuroscience* **46**, 2392–2405 (2017).

29. Chen, M. S. *et al.* Nogo-A is a myelin-associated neurite outgrowth inhibitor and an antigen for monoclonal antibody IN-1. *Nature* **403**, 434–439 (2000).

30. GrandPré, T., Nakamura, F., Vartanian, T. & Strittmatter, S. M. Identification of the Nogo inhibitor of axon regeneration as a Reticulon protein. *Nature* **403**, 439–444 (2000).

31. McKerracher, L. *et al.* Identification of myelin-associated glycoprotein as a major myelin-derived inhibitor of neurite growth. *Neuron* **13**, 805–811 (1994).

32. Wang, K. C. *et al.* Oligodendrocyte-myelin glycoprotein is a Nogo receptor ligand that inhibits neurite outgrowth. *Nature* **417**, 941–944 (2002).

33. Huang, J. K. *et al.* Glial membranes at the node of Ranvier prevent neurite outgrowth. *Science* **310**, 1813–1817 (2005).

34. Song, I. & Huganir, R. L. Regulation of AMPA receptors during synaptic plasticity. *Trends in Neurosciences* **25**, 578–588 (2002).

35. Barres, B. A. & Raff, M. C. Proliferation of oligodendrocyte precursor cells depends on electrical activity in axons. *Nature* **361**, 258–260 (1993).

36. Ishibashi, T. *et al.* Astrocytes Promote Myelination in Response to Electrical

Impulses. *Neuron* **49**, 823–832 (2006).

37. Han, M. *et al*. Individualized Cortical Parcellation Based on Diffusion MRI Tractography. *Cerebral Cortex* **30**, 3198–3208 (2020).
38. Kong, R. *et al*. Spatial Topography of Individual-Specific Cortical Networks Predicts Human Cognition, Personality, and Emotion. *Cerebral Cortex* **29**, 2533–2551 (2019).
39. Wang, D. *et al.* Parcellating cortical functional networks in individuals. *Nat Neurosci* **18**, 1853–1860 (2015).
40. Cieslak, M. *et al.* QSIPrep: an integrative platform for preprocessing and reconstructing diffusion MRI data. *Nat Methods* **18**, 775–778 (2021).
41. Jlr, A. & Sn, S. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage* **125**, (2016).
42. Yeh, F.-C. & Tseng, W.-Y. I. NTU-90: A high angular resolution brain atlas constructed by q-space diffeomorphic reconstruction. *NeuroImage* **58**, 91–99 (2011).
43. Garyfallidis, E. *et al*. Dipy, a library for the analysis of diffusion MRI data. *Frontiers in Neuroinformatics* **8**, (2014).
44. Avants, B. B., Epstein, C. L., Grossman, M. & Gee, J. C. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal* **12**, 26–41 (2008).
45. Jd, T. *et al.* MRtrix3: A fast, flexible and open software framework for medical

image processing and visualisation. *NeuroImage* **202**, (2019).

46. Fan, L. *et al.* The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture. *Cereb. Cortex* **26**, 3508–3526 (2016).
47. Chen, A. A. *et al.* Harmonizing functional connectivity reduces scanner effects in community detection. *NeuroImage* **256**, 119198 (2022).
48. Hawrylycz, M. J. *et al.* An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399 (2012).
49. Markello, R. D. & Arnatkevi, A. Standardizing workflows in imaging transcriptomics with the abagen toolbox. 23.
50. Arnatkeviciute, A., Markello, R. D., Fulcher, B., Misic, B. & Fornito, A. Towards best practices for imaging transcriptomics. Preprint at https://doi.org/10.31219/osf.io/y8ftn (2022).
51. Arnatkeviciute, A., Fulcher, B. D. & Fornito, A. A practical guide to linking brain-wide gene expression and neuroimaging data. *Neuroimage* **189**, 353–367 (2019).
52. Fulcher, B. D. & Fornito, A. A transcriptional signature of hub connectivity in the mouse connectome. *Proceedings of the National Academy of Sciences* **113**, 1435–1440 (2016).
53. Hawrylycz, M. *et al.* Canonical genetic signatures of the adult human brain. *Nat Neurosci* **18**, 1832–1844 (2015).
54. Wei, Y. *et al.* Statistical testing in transcriptomic□neuroimaging studies: A how□to and evaluation of methods assessing spatial and gene specificity. *Hum*

Brain Mapp **43**, 885–901 (2021).

55. Hansen, J. Y. *et al.* Mapping gene transcription and neurocognition across human neocortex. *Nat Hum Behav* **5**, 1240–1250 (2021).
56. Seidlitz, J. *et al.* Transcriptomic and cellular decoding of regional brain vulnerability to neurogenetic disorders. *Nat Commun* **11**, 3358 (2020).
57. Miller, J. A. *et al.* Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199–206 (2014).