Title: Mapping Consistent, Reproducible, and Transcriptionally Relevant Connectome Hubs of the Human Brain

Author affiliations: Zhilei Xu[a,b,c], Mingrui Xia [a,b,c], Xindi Wang [a,b,c], Xuhong Liao[d], Tengda Zhao [a,b,c], Yong He [a,b,c,e *]

[a] State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China

[b] Beijing Key Laboratory of Brain Imaging and Connectomics, Beijing Normal University, Beijing 100875, China

[c] IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China

[d] School of Systems Science, Beijing Normal University, Beijing 100875, China

[e] Chinese Institute for Brain Research, Beijing 102206, China

Running head: Consistency, reproducibility, and transcriptomic profiles of brain hubs

Corresponding author:

Yong He, Ph.D., E-mail: yong.he@bnu.edu.cn

Keywords: default mode; connectomics; development; neuroimaging; gene expression

Manuscript information: 28 text pages, 5 figures, 1 table.
(Supplementary Information: 22 text pages, 9 figures, 12 tables.)

Significance

The resting human brain includes some densely connected regions, known as functional connectome hubs. However, anatomical locations of these hubs are largely inconsistent and less reproducible among extant reports and the genetic signatures underlying hubs remain unknown. We identify highly consistent and reproducible hubs using a harmonized meta-connectomic analysis. These hubs are primarily located in heteromodal and unimodal regions and show unique, heterogeneous connectivity profiles. These hubs are also significantly associated with a transcriptomic pattern dominated by genes involved in the neuropeptide signaling pathway, neurodevelopmental processes, and cellular metabolic processes. These results highlight macroscopic hubs' robustness and their potential microstructural and metabolic substrates and have implications for understanding how brain hubs develop and function in health and disease.

Abstract

Macroscopic functional connectomic analyses have identified sets of densely connected regions in the human brain, known as connectome hubs, which play a vital role in understanding network communication, cognitive processing, and brain disorders. However, anatomical locations of functional connectome hubs are largely inconsistent and less reproducible among extant reports, partly due to inadequate sample size and differences in image processing and network analysis. Moreover, the genetic signatures underlying robust connectome hubs remain unknown. Here, we conduct the first worldwide voxelwise meta-connectomic analysis by pooling resting-state functional MRI data of 5,212 healthy young adults across 61 independent international cohorts with harmonized image processing and network analysis protocols. We identify highly consistent and reproducible functional connectome hubs that are spatially distributed in multiple heteromodal and unimodal regions, with the most robust findings mainly located in lateral parietal regions. These connectome hubs show unique, heterogeneous connectivity profiles and are critical for both intra- and inter-network communications. Using transcriptome data from the Allen Human Brain Atlas and BrainSpan Atlas as well as machine learning approaches, we demonstrate that these hubs are significantly associated with a transcriptomic pattern dominated by genes involved in the neuropeptide signaling pathway, neurodevelopmental processes, and cellular metabolic processes. This pattern represents microstructural and metabolic substrates underlying the development and functioning of brain hubs. Together, these results highlight the robustness of macroscopic connectome hubs of the human brain and their potential cellular and molecular underpinnings and have implications for understanding how brain hubs support the connectome organization in health and disease.

Introduction

Functional connectome mapping studies have identified sets of densely connected regions in large-scale human brain networks, which are known as hubs (1). Connectome hubs play a crucial role in global brain communication (1, 2) and support a broad range of cognitive processing, such as working memory (3, 4), semantic processing (5), choice reaction (6), and focused attention (7). Growing evidence suggests that these highly-connected brain hubs are preferentially targeted by many neuropsychiatric disorders (8-12), which provides critical clues for understanding the biological mechanisms of disorders and establishing biomarkers for disease diagnosis (10, 13) and treatment evaluation (14, 15) (reviewed in (1, 2, 16, 17)).

Despite such importance in global brain communication, cognitive processing, and brain disorders, there is considerable inconsistency in anatomical locations of functional connectome hubs among existing studies. For example, components of the default-mode network (DMN) have been frequently reported as connectome hubs, yet the spatial pattern is highly variable across studies. In particular, several studies have shown highly connected hubs in the lateral parietal regions of the DMN (8, 10, 18, 19), whereas others have reported midline structures of the DMN (20-23). Several works have identified primary sensorimotor and visual regions as connectome hubs (19-24), yet others did not replicate these findings (8, 10). Subcortical regions, such as the thalamus and amygdala, have also been inconsistently reported as hubs (10, 18, 21, 23) and non-hubs (8, 19, 20, 24). Sources of inconsistency in prior reports can be attributed to an inadequate sample size and differences in imaging scanner, imaging protocol, data preprocessing, and network analysis strategies. Thus, the consistency and reproducibility of functional connectome hubs have been difficult to establish to date. Here, we aimed to establish a harmonized meta-analytic model to identify robust functional connectome hubs in healthy young adults by combining multiple cohorts with standardized protocols for data quality assurance, image processing, and network analyses.

Once the robust connectome hubs are identified, we will further examine their associations with gene expression profiles. It has been well demonstrated that the functional connectome architecture of the human brain is inheritable, such as intranetwork (25, 26) and internetwork (27) connections and regional cost-efficiency balance (28). In addition, the brain's functional connectomes can be regulated by genotypic variation both during rest (29, 30) and in cognitive tasks (31, 32), especially involving the DMN (29, 32) and frontoparietal network (FPN) (30, 32). Moreover, transcriptome-connectome association investigations have provided evidence for spatial correspondence between transcriptomic profiles and functional connectome architectures (33-36), such as a strong nexus between the transcription of genes involved in potassium ion channel activity and the dynamic brain connectivity pattern (36). Thus, we reasoned that the robust connectome hubs of the macroscopic functional networks could be associated with microscopic genetic signatures. Elucidating these genetic signatures will benefit our understanding of how functional connectome hubs emerge in development, function in complex cognition, and are involved in disease.

To address these issues, we conducted the first worldwide, harmonized meta-connectomic analysis by pooling resting-state functional MRI (rsfMRI) data of 5,212 young adults across 61 independent cohorts. We identified highly consistent and reproducible connectome hubs in multiple heteromodal and unimodal regions, with the most robust findings occurring in the lateral parietal cortex. These connectome hubs were spatially distributed in multiple brain

networks, including the DMN, FPN, dorsal/ventral attention network (DAN/VAN), somatomotor network (SMN), and visual network (VIS), and showed heterogeneous connectivity profiles to provide support for both intra- and inter-network communications. Using microarray-based gene expression data from the Allen Human Brain Atlas (AHBA) (37) and BrainSpan Atlas (38), we demonstrated that these robust connectome hubs were linked with a transcriptomic pattern dominated by genes related to the neuropeptide signaling pathway, neurodevelopmental processes, and cellular metabolic processes.

Results

We collected a large-sample rsfMRI dataset (N = 6,830) from public data-sharing platforms, which consists of 61 independent international cohorts from Asia, Europe, North America, and Australia. After stringent quality controls, data from 5,212 healthy young adults ($18 \leq$ age < 36 years, 2,377 males) were included in our final analysis. Figure 1 illustrates the sample size and age ranges of each cohort. Details on participant inclusion criteria and quality control are provided in the *SI Appendix*, Sections 1.1 and 1.2.

Highly Consistent Connectome Hubs Using a Harmonized Meta-Analysis Model. Prior to the meta-analysis, we harmonized the image preprocessing and connectome construction and analysis across cohorts. Specifically, all rsfMRI data from each cohort were first preprocessed routinely with a standardized, uniform pipeline (*SI Appendix*, Section 1.2). Then, for each individual, we constructed a high-resolution functional connectome matrix by computing the Pearson's correlation coefficient between preprocessed rsfMRI time series of all pairs of gray matter voxels (N = 47,619). Negative functional connections were excluded from our analysis due to neurobiologically ambiguous interpretations (39). To further reduce signal noise and simultaneously avoid potential sharing signals between nearby voxels, both weak connections (Pearson's $r < 0.1$) and connections terminating within 20 mm were set to zero (40). Then, the functional connectivity strength (FCS) of each voxel was computed as the sum of connection weights between the given voxel and all the other voxels. This resultant FCS map was further normalized with respect to its mean and standard deviation across voxels (8). For each cohort, we performed a general linear model on these normalized FCS maps to reduce age and gender effects (*SI Appendix*, Section 1.3). As a result, we obtained a mean FCS map and a variance FCS map for each cohort that were used for subsequent meta-analyses.

To identify the most consistent connectome hubs, we conducted a voxelwise random-effects meta-analysis on the mean and variance FCS maps of 61 independent cohorts (*SI Appendix*, Section 1.4). Such an analysis addressed the across-cohort heterogeneity of functional connectomes, resulting in a robust FCS pattern and its corresponding variance (standard error, SE) map (Figure 2A). Furthermore, we computed voxelwise Z value by dividing the FCS map by its variance map. To determine the significance levels of these observed Z values, a nonparametric permutation test (41) with 10,000 iterations was performed. We identified connectome hub voxels (15,461 voxels, 32.5%) in which Z values were significantly ($p < 0.001$, cluster size > 200 mm^3) higher than the global mean (i.e., zero). Finally, we estimated voxelwise effect sizes using Cohen's d metric, which was computed by dividing the Z value map by the square root of the cohort number (Figure 2B, left). According to a brain network parcellation provided by (42, 43), these identified hub voxels were spatially distributed in multiple brain

networks, including the DMN (27.5%), DAN (16.5%), FPN (15.9%), VAN (15.6%), SMN (14.4%), and VIS (9.9%) (Figure 2B, right). Using a local maxima localization procedure, we further identified 35 robust connectome hub regions across 61 cohorts (Figure 2B, left), involving various heteromodal and unimodal areas. Specifically, the most robust findings resided in several lateral parietal regions, including the bilateral ventral postcentral gyrus, supramarginal gyrus, and angular gyrus. Table 1 illustrates the anatomical locations of these hubs and the MNI coordinates of their peak voxels.

Highly Reproducible Connectome Hubs Using Leave-One-Cohort-Out Validation Analysis and Conjunction Analysis. Despite the consistency of connectome hubs identified above, the random-effects meta-analysis revealed high heterogeneity of FCS across cohorts (Figure 2C, left). The cumulative distribution function plot shows more than 95% voxels with I^2 (heterogeneity score) exceeding 50% (Figure 2C, right), indicating high heterogeneity across cohorts in almost all brain areas. To determine whether the connectome hubs identified here are dominated by certain cohorts or are reproduced across-subject/cohort, we performed a leave-one-cohort-out validation analysis and an across-subject/cohort conjunction analysis. (i) *Leave-one-cohort-out validation analysis.* We repeated the above meta-analytic hub identification procedure after leaving one cohort out at a time. Comparing the identified hubs using all cohorts (Figure 2B) with those after leaving one cohort out obtained extremely high Dice's coefficients (*mean*±*sd*: 0.990±0.006; range: 0.966-0.997). For hub peaks, leaving one cohort out resulted in very few displacements (mostly fewer than 6 mm, Figure 2D). Thus, functional connectome hubs identified using the 61 cohorts were not dominated by specific cohorts. (ii) *Across-subject/cohort conjunction analysis.* We defined the top N (N = 15,461, which is the voxel number of hubs identified in Figure 2B) voxels with the highest FCS values of a subject or a cohort as connectome hubs for that subject or that cohort. Then, for each voxel, we assessed hub occurrence probability (HOP) values across subjects and cohorts. The identified hubs using all cohorts were highly overlapped with the top N voxels with the highest HOP values both across all subjects and across all cohorts, indicated by a high Dice's coefficient (*Dice*=0.867, Figure 2E, left; *Dice*=0.924, Figure 2E, right). When the connectome hubs from all cohorts were compared with the top N voxels with the highest HOP values across randomly selected subjects or across randomly selected cohorts, the Dice's coefficient approached 99% of its maximum value after exceeding 510 subjects (Figure 2F, left) and 35 cohorts (Figure 2F, right), respectively. Validation analysis demonstrated that the above results did not depend on analysis parameters, such as the connection threshold (Figures S1 and S2). Together, these results suggest that rsfMRI data of 5,212 subjects from 61 cohorts are sufficient to map highly consistent and reproducible functional connectome hubs in the resting human brain.

Heterogeneous Connectivity Profiles of Connectome Hubs and Their Distinctive Roles in Bridging Brain Networks. Next, we further examined whether these robust brain hubs identified here (Figure 2B and Table 1) have distinctive connectivity profiles that represent their unique roles in brain network communication. To gain detailed and robust functional connectivity profiles of each hub region, we conducted a seed-to-whole-brain connectivity meta-analysis in a harmonized protocol again (*SI Appendix*, Section 1.5). For each of the 35 hub regions, we obtained an estimated Cohen's d effect size map that characterizes the robust whole-brain connectivity patterns functionally relevant to the seed region across 61 independent cohorts (Figure 3A). We then divided the connectivity map of each hub into eight brain networks according to a brain network parcellation provided by (42, 43) and represented the functional

connectivity profile of a hub as the voxel percentage of each of the eight networks connected with it to address the effect of network size. As a result, we obtained an 8×35 percentage matrix, with each column representing the robust connectivity profile of a hub.

Hierarchical clustering analysis on the connectivity percentage matrix clearly divided the 35 hub regions into three clusters (Figure 3B). Cluster I consists of 21 hubs that are primarily connected with extensive areas in the DAN, VAN, FPN, and SMN (orange color, Figure 3B). Cluster II consists of four hubs that are densely connected with VIS (green color, Figure 3B). Cluster III consists of 10 hubs that have robust connections with the DMN and LIMB (blue color, Figure 3B). Of particular interest is that within Cluster III, a left posterior middle frontal hub called ventral area 8A (8Av) shows a distinctive connectivity profile in contrast to the other nine hubs, manifested as having no functional connections with the hippocampal-entorhinal complex but having robust functional connections with bilateral lateral frontal FPN regions (Figure 3A). These results imply that the left 8Av hub is a key connector between the DMN and FPN, which can be supported by the recent finding of a control-default connector located in the posterior middle frontal gyrus (44).

Thus, whereas all hubs possess dense intranetwork functional connections, most also retain significant internetwork connections to bridge large-scale brain networks, which preserves efficient communication across the whole brain network feasible. More details can be found in the *SI Appendix*, Section 2.1 and Figures S3 and S4.

Transcriptomic Signatures Associated with the Robust Connectome Hubs. Considering the significant heritability of human brain functional connectivity (25-32) and substantial evidence supporting genetic factors involved in both the synchronous activity (33) and the spatial organization (34-36) of the human brain network, we reasoned that significant genetic signatures are associated with the robust functional connectome hubs identified here.

Distinguishing connectome hubs using machine learning approaches. To address this issue, we used a machine learning approach to examine whether connectome hubs can be distinguished from non-hubs based on their transcriptomic profiles. We extracted 10,027 genes' preprocessed transcriptomic data (45) from the AHBA (37) for 1,158 left cortical samples, including 382 hub samples and 776 non-hub samples (Figure 4A, Table S3). A supervised machine learning classifier based on XGBoost (46) was trained with 300 randomly selected hub samples and 300 randomly selected non-hub samples to distinguish hub samples from non-hub samples using the 10,027 genes' transcriptomic features and was tested with the remaining 82 hub samples and 476 non-hub samples. Before training the classifier, we determined the optimal training iterations through a cross-validation procedure. The sensitivity, specificity, and accuracy rate of the XGBoost classifier were stably estimated by repeating the randomly selecting training samples, cross-validation, classifier training, and testing procedure 1,000 times.

The XGBoost classifier performed better than chance in all 1,000 repetitions and achieved an overall accuracy rate of 65.3% (Figure 4B). In cross-validation, hub samples and non-hub samples were classified correctly with a sensitivity of 71.1% and specificity of 63.4%, respectively. The testing procedure yielded a comparable sensitivity of 69.7% and specificity of 62.0%. As expected, brain samples with greater absolute Cohen's *d* values possessed a higher rate to be correctly classified ($ps < 0.0001$, 10,000 permutation tests, Figure 4C). After training the classifier, each gene's contribution to the optimal prediction model was determined. We noted that some key genes contributed two or three orders of magnitude more than other genes

(Figure 4D, Table S8). Importantly, the contributions of the top 300 mostly contributed key genes were consistent between the first 500 repetitions and the second 500 repetitions (Pearson's $r = 0.958$, $p < 10^{-6}$, Figure 4E), which demonstrated high reproducibility.

To exclude the XGBoost model's potential bias relating to the mostly contributed key genes, we replicated the classification results above using another machine learning model based on the support vector machine (SVM) (47, 48). As illustrated in Figure 4F, we first trained an SVM classifier through a cross-validation procedure using the transcriptomic features of the top *N* key genes with the greatest contribution to the XGBoost classifier. Because no data were available to determine how many key genes were sufficient to train an SVM classifier, we examined the count from 100 to 300. The SVM classifier achieved a very high peak accuracy rate of 91.8% with approximately the top 150 key genes in the easiest classification task to distinguish all 382 hub samples from 382 non-hub samples with the highest rate to be correctly classified by the XGBoost classifier (Figure 4G). The SVM classifier also achieved a reasonable peak accuracy rate of 67.8% with approximately the top 150 key genes even in the most difficult classification task to distinguish all 382 hub samples from 382 non-hub samples with the lowest rate to be correctly classified by the XGBoost classifier (Figure 4H and 4I). However, SVM classifiers trained using 150 randomly selected genes performed significantly worse than that using the top 150 key genes (*ps* < 0.001, Figure 4I). Thus, these robust functional connectome hubs were significantly associated with a transcriptomic pattern dominated by approximately 150 key genes.

Gene ontology enrichment analysis. Gene Ontology (GO) enrichment analysis using GOrilla (http://cbl-gorilla.cs.technion.ac.il/, (49)) demonstrated that these 150 key genes were mostly enriched in the neuropeptide signaling pathway (*fold enrichment* (*FE*) = 8.9, $p = 1.2\times10^{-5}$, Table S9). GO enrichment analysis using all the 10,027 genes' contributions to the XGBoost classifier also confirmed the most significant GO term of the neuropeptide signaling pathway (*FE* = 5.7, $p = 0.3\times10^{-7}$, Table S10). The contributions of these 10,027 genes were also significantly associated with the developmental process (*FE* = 1.2), cellular developmental process (*FE* = 1.3), cell differentiation (*FE* = 1.4), and neuron projection arborization (*FE* = 13.7) (*ps* < 0.0006, Table S10). Moreover, the HSPB8 (heat shock protein family B (Small) member 8) gene, whose transcription is significantly associated with dendritogenesis and myelination (50), emerged as one of the top 150 key genes. These results suggest that functional connectome hubs may experience distinctive neurodevelopmental processes in contrast to non-hubs. To gain more details, we further examined regional transcriptomic differences between hub and non-hub regions for genes previously implicated in key neurodevelopmental processes (50) (Table S5). Permutation tests revealed hub regions with significantly higher transcription levels for genes associated with dendrite ($p = 0.0045$) and synapse ($p = 0.0030$) development than non-hub regions (Figure 5A). In addition, hub regions had significantly lower transcription levels for genes associated with axon development ($p = 0.0318$), myelination ($p = 0.0386$), and neuron migration ($p = 0.0132$) but not for neuron differentiation ($p = 0.4155$) (Figure 5A).

These above transcriptomic results were derived from the AHBA, a healthy adult transcriptomic dataset. To obtain temporal details, we inspected regional transcriptomic trajectory differences in these key neurodevelopmental processes using the BrainSpan Atlas (38). We observed pronounced diverging transcriptomic trajectories between hub and non-hub regions for genes associated with neuron migration, dendrite, synapse, axon development, and myelination but not for neuron differentiation (Figure 5B). For neuron migration, the transcription level in hub

regions is higher than that in non-hub regions since the mid-fetal period and until after birth. For dendrite and synapse development, transcriptomic trajectories of hub regions diverge from those of non-hubs since early childhood and until the end of the dataset at 40 years, during which hub regions have higher transcription levels for most periods. Diverging emerges earlier for axon development and myelination since the late fetal period and diminishes at the end of adolescence. Conversely, higher transcription levels are nearly always observed in non-hub regions. These results are in agreement with the observation of primary somatosensory, auditory, and visual (V1/V2) cortices with lower synapse density but higher myelination than the prefrontal area in (51, 52). These microscale divergences of key neurodevelopmental processes may result in distinctive macroscale connectivity patterns in hub regions in contrast to non-hubs. Indeed, using a fiber length profiling dataset provided by (53), we observed more fibers with a length exceeding 40 mm in hub regions than in non-hubs, indicating hub regions possessing more short, medium, and long fibers, whereas non-hub regions possessed more very short (< 40 mm) fibers (Figure 5C).

To exclude the GO tool's potential bias relating to enrichment analysis results, we repeated the GO enrichment analysis of the top 150 key genes using DAVID (https://david.ncifcrf.gov/, (54, 55)). In addition to the most significant GO term of the neuropeptide signaling pathway (FE = 8.7, $p < 0.0006$), there were 10 GO terms associated with metabolic process, such as the positive regulation of cellular metabolic process (34 genes, $FE = 1.4$, $p = 0.031$) (Table S11). Disease association analysis demonstrated metabolic disease associated with the greatest number of key genes at a marginally significant level (60 genes, $FE = 1.2$, $p = 0.094$) (Table S12). Based on these two results, it is rational to speculate that functional connectome hubs possess distinctive metabolic patterns in contrast to non-hubs. To confirm this hypothesis, we further examined regional differences in main metabolic pathways in the human brain: oxidative phosphorylation and aerobic glycolysis (56). Oxidative phosphorylation and aerobic glycolysis were imaged using positron emission tomography and were estimated by the cerebral metabolic rate for oxygen ($CMRO_2$) and the glycolytic index, respectively (57). We found hub regions with significantly higher levels of both $CMRO_2$ ($p = 0.0038$) and glycolytic index ($p = 0.0004$) than non-hub regions (10,000 permutation tests, Figure 5D). We also inspected transcriptomic trajectories of genes associated with aerobic glycolysis in the human brain (58) using the BrainSpan Atlas (38). Consistent with the positron emission tomography results, aerobic glycolysis-associated genes showed higher transcription levels in hub regions than in non-hub regions since the late fetal period (Figure 5E). Adult transcriptomic data from the AHBA also confirmed this significant difference ($p < 0.0001$, 10,000 permutation tests). Of note, considering apparent transcriptomic differences compared to the neocortex (50), we excluded the striatum, mediodorsal nucleus of the thalamus, and cerebellar cortex in the transcriptomic trajectory analysis but not the amygdala and hippocampus, whose transcriptomic trajectories are more similar to those of the neocortex than to those of other subcortical structures (50, 59). We also conducted transcriptomic trajectory analysis using only neocortical regions and observed almost unchanged results (Figure S5).

Taken together, the robust functional connectome hubs identified here are significantly associated with a transcriptomic pattern dominated by genes related to the neuropeptide signaling pathway, neurodevelopmental processes, and cellular metabolic processes.

Discussion

Using a harmonized meta-connectomic analysis of 5,212 young adults across 61 independent cohorts, we identified highly consistent and reproducible functional connectome hubs in multiple heteromodal and unimodal regions. Using transcriptome data from the AHBA and BrainSpan Atlas, we demonstrated distinctive spatiotemporal transcriptomic signatures underlying functional connectome hubs. These results markedly further our understanding of the robustness of macroscopic connectome hubs of the human brain and their potential microscopic cellular and molecular substrates.

Challenges for Identifying Highly Consistent and Reproducible Functional Connectome Hubs. Extant reports have shown largely inconsistent and less reproducible hub localizations (8, 10, 18-24), which may arise from high heterogeneity in the included subjects, data acquisition, and analysis methods across studies. To diminish these potential confounding factors, we employed stringent participant inclusion criteria that included only healthy young adults aged 18 to 36 years in the final analysis, and adopted harmonized data preprocessing and connectome analysis protocols across cohorts. Nevertheless, random-effects meta-analysis revealed high heterogeneity among cohorts in almost all brain areas. For more than 95% gray matter voxels, the variation of FCS across subjects was dominated by heterogeneity among cohorts. Combined with previous findings of apparent site-related variation of functional connectivity throughout the brain (60), we speculate that high heterogeneity among cohorts caused by imaging scanners and/or imaging protocols is an important cause for inconsistent and less reproducible hub localizations among prior studies. Thus, combining data from multiple cohorts and adopting a harmonized random-effects meta-analysis model were both indispensable in the present study.

High heterogeneity among cohorts raised the concern of whether the hubs identified here were dominated by certain cohorts. We examined it by repeating the functional connectome hub identification procedure after leaving one cohort out at a time, and demonstrated that the hubs identified using the 61 cohorts were not dominated by any cohort. This may benefit from our rational statistical model and sufficient datasets. We adopted a harmonized random-effects meta-analysis model to handle both intracohort variation (i.e., sampling errors) and intercohort heterogeneity (61). Our results showed that the spatial distribution of functional connectome hubs was relatively stable when using more than 510 subjects and 35 cohorts, demonstrating that 5,212 subjects from 61 cohorts were adequate to minimize both sampling errors and heterogeneity among cohorts. It also provided guidelines for mapping robust functional connectivity of the human brain across populations in future studies. Considering that there were only dozens of subjects in most prior studies (8, 10, 18-20, 22, 24), we speculate that low statistical power attributed to inadequate subjects is another cause for prior inconsistent and less reproducible hub localizations (62). It is worthy to note that the consistent functional connectome hubs identified with harmonized image processing and network analysis were highly reproducible both across subjects and across cohorts. This suggests that methodological variation or defects may also cause controversial hub reports in specific regions across prior studies.

Together, the inconsistency and less reproducibility of hub localizations reported previously may be mainly due to high heterogeneity among cohorts in the included subjects, data acquisition, and analysis methods and low statistical power attributed to inadequate subjects. These potential confounding factors were thoroughly addressed in the present study through stringent participant inclusion criteria, a rational harmonized random-effects meta-analysis model, and the

combination of adequate subjects from multiple cohorts. An extension of discussion can be found in the *SI Appendix*, Section 3.1 and Figures S6 to S8.

The Highly Consistent and Reproducible Functional Connectome Hubs. Prior studies have reported functional connectome hubs throughout heteromodal, unimodal, primary, and subcortical regions, but the results are inconsistent and less reproducible across studies (8, 10, 18-24). Here, we demonstrated highly consistent and reproducible functional connectome hubs mainly located in the DMN, DAN, VAN, FPN, SMN, and VIS. Using a local maxima localization procedure, we further identified 35 robust connectome hubs whose distinctive functional connectivity profiles clearly divided them into three clusters. These functional connectivity profiles provided intriguing clues about their distinctive roles in brain network communication.

Twenty-one hubs (Cluster I) are robustly connected with extensive areas in the DAN, VAN, FPN, and SMN, indicting their favoring global integration of information processing among these networks. In fact, they are exactly core regions forming the DAN, VAN, and FPN as well as the sensorimotor pathway. Specifically, four lateral frontal hubs left 43, left FOP4, right area 46, and right 6r, two ventral postcentral hubs right PF and left PFop, and two medial hubs left SCEF and left 5mv are canonical VAN regions reported in (42, 63), whereas four lateral parietal hubs left AIP, right 7PC, and bilateral PFt, two dorsal frontal hubs bilateral 6a, one ventral frontal hub right 6v, and one temporal hub FST are core DAN regions (63). These VAN/DAN hub regions' differential responses to goal-directed and stimulus-driven attention tasks (64) also clearly manifested as nuanced functional connectivity profiles in the present study although with many overlapping regions. In addition, two lateral prefrontal hubs left p9-46v and right IFSa are core FPN regions that have been reported both in group-level seed-based (65) and in individual-level connectome-based (66) investigations. Moreover, hub regions involved in the sensorimotor pathway act as significant waystations for communication between the VIS and the SMN, DAN, and VAN, such as the right VIP, right FST, left 7Am, and left FEF, which are robustly connected with the visual association cortex, such as the MT+ complex and neighboring visual areas and peripheral visual areas. Information flow along the primary visual, visual association, and higher-level sensorimotor cortices is undertaken by the four occipital hubs (Cluster II) left VMV1, right V4, and bilateral V3A, which are all densely connected with the VIS and portion of the SMN regions. This aligns with the role of their homologous regions in the non-human primate cerebral cortex, which has been validated by seminal tract tracing work (67).

The remaining 10 hubs (Cluster III) are all located in canonical DMN regions (68). Nine of them confine functional connectivity within the DMN and LIMB regions, whereas the left 8Av hub acts as a connector between the DMN and FPN that manifests as being robustly connected with both DMN regions and lateral prefrontal FPN regions. This can be supported by the recent finding of a control-default connector located in the posterior middle frontal gyrus (44) and may also be a case of the hypothesis of a parallel interdigitated subnetwork (69). These observations offer a significant complementary interpretation to the conventional assumption that the DMN is anticorrelated with other networks (68, 70). Additionally, considering that communication between the DMN and other networks is of significant relevance to neuropsychiatric disorders, such as major depressive disorder (71) and autism spectrum disorders (72), the left 8Av hub may be a promising target region for therapeutic interventions.

Spatiotemporal Transcriptomic Signatures Underlie the Robust Functional Connectome Hubs. To our knowledge, this is the first study to demonstrate spatiotemporal transcriptomic signatures underlying the robust connectome hubs in the human brain. Unsurprisingly, these brain hubs were significantly associated with a transcriptomic pattern dominated by genes with the highest enrichment for the neuropeptide signaling pathway. Because neuropeptides are a main type of synaptic transmitter that is widely distributed in the human central nervous system (73), the neuropeptide signaling pathway is indispensable for efficient synaptic signal transduction that sustains dense and flexible functional connections in hub regions. In addition, 34 and 60 out of the 150 key genes were involved in the positive regulation of cellular metabolic process and metabolic disease, respectively, which suggests that hub regions may possess distinctive metabolic patterns in contrast to non-hubs. This is possible and reasonable because massive synaptic activities in hub regions demand high material and metabolic costs, which is in accordance with our observation of higher oxidative phosphorylation and aerobic glycolysis levels in hub regions. This is also compatible with prior findings of higher cerebral blood flow (74) and aerobic glycolysis levels (75) in hub regions.

We found functional connectome hubs possessing a distinctive transcriptomic pattern of key neurodevelopmental processes in contrast to non-hubs. Compared with non-hub regions, the connectome hubs have higher transcription levels of genes associated with dendrite and synapse development but lower levels of axon development and myelination during adulthood. These differences are also apparent during development, especially during childhood and adolescence. Our findings are compatible with previous observations of the prefrontal area having higher synapse density but lower myelination than primary somatosensory, auditory, and visual (V1/V2) cortices (51, 52). Importantly, differences observed in (51) are also more apparent during childhood and adolescence. Higher transcription levels for dendrite and synapse development in hub regions are necessary for the overproduction of synapses that will be selectively eliminated based on the demand of the environment and gradually stabilized before full maturation (76), which is “the major mechanism for generating diversity of neuronal connections beyond their genetic determination” (76, 77). Lower transcription levels for axon development and myelination will prolong the myelination period in hub regions, which characterizes a delayed maturation phase (78). Marked delay of anatomical maturation in human prefrontal and lateral parietal cortices has been frequently observed both in human development (77, 79, 80) and in primate evolution (78), which provides more opportunities for social learning to establish diverse neuronal circuits that contribute to our complex (77) and species-specific (78) cognitive capabilities. We also observed higher transcription levels for neuron migration in hub regions during most fetal periods and early infancy. This is in agreement with the report of extensive migration of young neurons persisting for several months after birth in the human frontal cortex (81). Meanwhile, the migration and final laminar positioning of postmitotic neurons are regulated by common transcription factors (82), which suggests that a higher transcription level for neuron migration in hub regions may facilitate the construction of more intricate interlaminar connectivity. These microscale divergences of key neurodevelopmental processes may result in distinctive macroscale anatomical connectivity patterns in hub regions in contrast to non-hubs, because we observed more short, medium, and long fibers in hub regions than in non-hubs. Consequently, this distinctive transcriptomic pattern may underlie the development of intricate neuronal circuits in hub regions that provide anatomical substrates for our complex cognitive capabilities later in life.

Human neurodevelopment is an intricate and protracted process, during which the transcriptome of the human brain requires precise spatiotemporal regulation (83). Thus, in addition to contributing to our complex cognitive capabilities, the distinctive transcriptomic pattern of neurodevelopment in hub regions may also enhance our susceptibility to neuropsychiatric disorders, which means small disturbance in the magnitude or the timing of this distinctive transcriptomic pattern may have long-term consequences on brain anatomical topography or functional activation. This is in line with our observation of psychiatric disorders being the most significant disease associated with the top 150 key genes (Table S12). These results imply that uncovering the intricate transcriptomic pattern, diverse neuronal circuits, anatomical topography, and functional activation of hub regions provide crucial and promising routes for understanding the pathophysiological mechanisms underlying neurodevelopmental disorders, such as autism spectrum disorders (11, 72), attention deficit hyperactivity disorder (11), and schizophrenia (9).

Of note, we conducted a data-driven analysis using machine learning based on XGBoost and SVM that were implemented through non-linear mathematical operations rather than linear operations, such as linear correlation (33), linear regression (84), principal component analysis (85), or partial least squares (35, 36). It has been argued that observations of transcriptomic-neuroimaging association through linear regression have a false-positive rate as high as 60% although with a stringent permutation test (86). Moreover, the results derived by principal component analysis and partial least squares may be largely shifted toward the first principal component axis of the dataset (87). These reports suggest that prior results of transcriptomic-neuroimaging association using the AHBA derived by linear mathematical operations may include highly reproducible false-positive observations that are independent of neuroimaging measurements, such as the association with genes enriched for ion channels (33, 36, 84, 85) and mitochondria (35, 36). By contrast, high reproducibility across different machine learning models and across different GO enrichment analysis tools and convergent results from the AHBA, BrainSpan Atlas, and multimodal neuroimaging data from other independent laboratories made it very unlikely that our findings were false-positive observations.

Methodological Considerations. Some results of the present study should be interpreted cautiously because of methodological issues. First, we were unable to resolve whether functional connectome hubs in the superior temporal gyrus and rolandic operculum were derived by an artifact of unavoidable signal blurring due to rsfMRI preprocessing or close proximity, which should be verified in future studies using rsfMRI with higher spatial resolution and greater signal specificity. Second, these above robust connectome hubs were identified using preprocessed rsfMRI data with global signal regression. We also repeated identifying functional connectome hubs using preprocessed rsfMRI data without global signal regression and observed a shifted hub distribution, which, however, may be derived from physiological artifacts rather than by the brain's intrinsic or ongoing neural activity (*SI Appendix*, Section 2.2 and Figure S9). Third, the AHBA dataset only includes partial human genes, of which approximately half were excluded in data preprocessing (45), which may have induced incomplete observations in our data-driven analysis. Finally, our transcriptomic signature results addressed only the association between functional connectome hubs and transcriptomic patterns and did not explore causation between them. Exploring more detailed mechanisms underlying this correlation is attractive and may be practicable for non-human primate brains in future studies.

Materials and Methods

The rsfMRI data of 6,830 subjects were collected from public data-sharing platforms and were preprocessed routinely with a standardized, uniform pipeline. Preprocessed rsfMRI data of 5,212 subjects from 61 cohorts were included in our final analysis. Harmonized random-effects meta-analyses were performed to identify robust connectome hubs and their connectivity profiles. The AHBA and BrainSpan Atlas as well as machine learning models based on XGBoost and SVM were used to identify transcriptomic signatures associated with robust connectome hubs. Additional information regarding methods is available in the *SI Appendix*, Section 1 and Tables S1 to S6. Code and data that support the findings of this study are publicly available at https://github.com/zhileixu/FunctionalConnectomeHubs.

Acknowledgements

We thank the International Neuroimaging Data-sharing Initiative, Brain Genomics Superstruct Project (88), Human Connectome Project, MPI-Leipzig Mind-Brain-Body Project, and Age-ility Project for providing publicly available neuroimaging datasets, and thank the Allen Institute for Brain Science for providing the AHBA and BrainSpan Atlas. We thank Drs. Huali Wang and Xiaodan Chen for data acquisition of the PKU cohort, and Drs. Qiushi Wang and Nan Zhang for advising on data quality assurance. This study was supported by the National Natural Science Foundation of China (Nos. 82021004, 31830034, 81620108016, 82071998, 81671767, 81971690, and 81801783), Changjiang Scholar Professorship Award (No. T2015027), National Key Research and Development Project (No. 2018YFA0701402), Beijing Nova Program (No. Z191100001119023), and Fundamental Research Funds for the Central Universities (No. 2020NTST29).

Competing Interests

Authors declare that they have no competing interests.

Author contributions:

Z.X. and Y.H. designed research; Z.X. performed research; Z.X. analyzed data; M.X., X.W., X.L., and T.Z. advised on research design; M.X. and X.W. advised on data analysis; M.X., X.L., T.Z., and X.W. advised on paper writing; and Z.X. and Y.H. wrote the paper.

References

1. M. P. van den Heuvel, O. Sporns, Network hubs in the human brain. *Trends Cogn Sci* **17**, 683-696 (2013).
2. E. Bullmore, O. Sporns, The economy of brain network organization. *Nat Rev Neurosci* **13**, 336-349 (2012).
3. M. W. Cole, T. Yarkoni, G. Repovs, A. Anticevic, T. S. Braver, Global connectivity of prefrontal cortex predicts cognitive control and intelligence. *J Neurosci* **32**, 8988-8999 (2012).
4. J. Liu *et al.*, Intrinsic Brain Hub Connectivity Underlies Individual Differences in Spatial Working Memory. *Cerebral Cortex* **27**, 5496-5508 (2016).
5. Y. Xu, Q. Lin, Z. Han, Y. He, Y. Bi, Intrinsic functional network architecture of human semantic processing: Modules and hubs. *Neuroimage* **132**, 542-555 (2016).
6. R. Leech, R. Braga, D. J. Sharp, Echoes of the brain within the posterior cingulate cortex. *J Neurosci* **32**, 215-222 (2012).
7. K. A. Garrison *et al.*, Real-time fMRI links subjective experience with brain activity during focused attention. *Neuroimage* **81**, 110-118 (2013).
8. R. L. Buckner *et al.*, Cortical Hubs Revealed by Intrinsic Functional Connectivity: Mapping, Assessment of Stability, and Relation to Alzheimer's Disease. *The Journal of Neuroscience* **29**, 1860 (2009).
9. M. P. van den Heuvel *et al.*, Abnormal rich club organization and functional brain dynamics in schizophrenia. *JAMA Psychiatry* **70**, 783-792 (2013).
10. Z. Dai *et al.*, Identifying and Mapping Connectivity Patterns of Brain Network Hubs in Alzheimer's Disease. *Cerebral Cortex* **25**, 3723-3742 (2014).
11. S. Ray *et al.*, Structural and functional connectivity of the human brain in autism spectrum disorders and attention-deficit/hyperactivity disorder: A rich club-organization study. *Hum Brain Mapp* **35**, 6032-6048 (2014).
12. N. A. Crossley *et al.*, The hubs of the human connectome are generally implicated in the anatomy of brain disorders. *Brain* **137**, 2382-2395 (2014).
13. J. Wang *et al.*, Disrupted functional brain connectome in individuals at risk for Alzheimer's disease. *Biol Psychiatry* **73**, 472-481 (2013).
14. L. Wang *et al.*, The effects of antidepressant treatment on resting-state functional brain networks in patients with major depressive disorder. *Hum Brain Mapp* **36**, 768-778 (2015).
15. J. S. Perrin *et al.*, Electroconvulsive therapy reduces frontal cortical connectivity in severe depressive disorder. *Proceedings of the National Academy of Sciences* **109**, 5464 (2012).
16. A. Fornito, A. Zalesky, M. Breakspear, The connectomics of brain disorders. *Nat Rev Neurosci* **16**, 159-172 (2015).
17. Q. Gong, Y. He, Depression, Neuroimaging and Connectomics: A Selective Overview. *Biological Psychiatry* **77**, 223-235 (2015).
18. M. W. Cole, S. Pathak, W. Schneider, Identifying the brain's most globally connected regions. *Neuroimage* **49**, 3132-3148 (2010).
19. X.-H. Liao *et al.*, Functional brain hubs and their test–retest reliability: A multiband resting-state functional MRI study. *NeuroImage* **83**, 969-982 (2013).
20. S. Achard, R. Salvador, B. Whitcher, J. Suckling, E. Bullmore, A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *J Neurosci* **26**, 63-72 (2006).
21. D. Tomasi, N. D. Volkow, Functional connectivity density mapping. *Proceedings of the National Academy of Sciences* **107**, 9885-9890 (2010).

22. P. Fransson, U. Aden, M. Blennow, H. Lagercrantz, The functional architecture of the infant brain as revealed by resting-state fMRI. *Cereb Cortex* **21**, 145-154 (2011).
23. D. Tomasi, N. D. Volkow, Association between Functional Connectivity Hubs and Brain Networks. *Cerebral Cortex* **21**, 2003-2013 (2011).
24. F. de Pasquale *et al.*, The connectivity of functional cores reveals different degrees of segregation and integration in the brain at rest. *Neuroimage* **69**, 51-61 (2013).
25. D. C. Glahn *et al.*, Genetic control over the resting brain. *Proc Natl Acad Sci U S A* **107**, 1223-1228 (2010).
26. J. Xu *et al.*, Heritability of the Effective Connectivity in the Resting-State Default Mode Network. *Cereb Cortex* **27**, 5626-5634 (2017).
27. Z. Yang *et al.*, Genetic and Environmental Contributions to Functional Connectivity Architecture of the Human Brain. *Cereb Cortex* **26**, 2341-2352 (2016).
28. A. Fornito *et al.*, Genetic influences on cost-efficient organization of human cortical functional networks. *J Neurosci* **31**, 3261-3270 (2011).
29. J. L. Wiggins *et al.*, The impact of serotonin transporter (5-HTTLPR) genotype on the development of resting-state functional connectivity in children and adolescents: a preliminary report. *Neuroimage* **59**, 2760-2770 (2012).
30. H. Long *et al.*, Interaction effect between 5-HTTLPR and HTR1A rs6295 polymorphisms on the frontoparietal network. *Neuroscience* **362**, 239-247 (2017).
31. C. Giessing, T. Neber, C. M. Thiel, Genetic variation in nicotinic receptors affects brain networks involved in reorienting attention. *Neuroimage* **59**, 831-839 (2012).
32. E. M. Gordon, M. Stollstorff, J. M. Devaney, S. Bean, C. J. Vaidya, Effect of dopamine transporter genotype on intrinsic functional connectivity depends on cognitive state. *Cereb Cortex* **22**, 2182-2196 (2012).
33. J. Richiardi *et al.*, BRAIN NETWORKS. Correlated gene expression supports synchronous activity in brain networks. *Science* **348**, 1241-1244 (2015).
34. F. M. Krienen, B. T. Yeo, T. Ge, R. L. Buckner, C. C. Sherwood, Transcriptional profiles of supragranular-enriched genes associate with corticocortical network architecture in the human brain. *Proc Natl Acad Sci U S A* **113**, E469-478 (2016).
35. P. E. Vertes *et al.*, Gene transcription profiles associated with inter-modular hubs and connection distance in human functional magnetic resonance imaging networks. *Philos Trans R Soc Lond B Biol Sci* **371** (2016).
36. J. Liu, M. Xia, X. Wang, X. Liao, Y. He, The spatial organization of the chronnectome associates with cortical hierarchy and transcriptional profiles in the human brain. *NeuroImage* **222**, 117296 (2020).
37. M. J. Hawrylycz *et al.*, An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391-399 (2012).
38. J. A. Miller *et al.*, Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199-206 (2014).
39. A. J. Schwarz, J. McGonigle, Negative edges and soft thresholding in complex network analysis of resting state functional connectivity data. *Neuroimage* **55**, 1132-1146 (2011).
40. Jonathan D. Power *et al.*, Functional Network Organization of the Human Brain. *Neuron* **72**, 665-678 (2011).
41. T. E. Nichols, A. P. Holmes, Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum Brain Mapp* **15**, 1-25 (2002).
42. B. T. Yeo *et al.*, The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* **106**, 1125-1165 (2011).

43. N. Tzourio-Mazoyer *et al.*, Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* **15**, 273-289 (2002).
44. E. M. Gordon *et al.*, Three Distinct Sets of Connector Hubs Integrate Human Brain Function. *Cell reports* **24**, 1687-1695 e1684 (2018).
45. A. Arnatkevic Iute, B. D. Fulcher, A. Fornito, A practical guide to linking brain-wide gene expression and neuroimaging data. *Neuroimage* **189**, 353-367 (2019).
46. T. Chen, C. Guestrin, XGBoost: A scalable tree boosting system. *arXiv:1603.02754* 10.1145/2939672.2939785 (2016).
47. C.-C. Chang, C.-J. Lin, LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2**, Article 27 (2011).
48. F. Pedregosa *et al.*, Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
49. E. Eden, R. Navon, I. Steinfeld, D. Lipson, Z. Yakhini, GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC bioinformatics* **10**, 48 (2009).
50. H. J. Kang *et al.*, Spatio-temporal transcriptome of the human brain. *Nature* **478**, 483-489 (2011).
51. C. C. Sherwood, A. Gómez-Robles, Brain Plasticity and Human Evolution. *Annual Review of Anthropology* **46**, 399-419 (2017).
52. P. R. Huttenlocher, A. S. Dabholkar, Regional differences in synaptogenesis in human cerebral cortex. *The Journal of comparative neurology* **387**, 167-178 (1997).
53. C. J. Bajada, J. Schreiber, S. Caspers, Fiber length profiling: A novel approach to structural brain organization. *Neuroimage* **186**, 164-173 (2019).
54. W. Huang da, B. T. Sherman, R. A. Lempicki, Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37**, 1-13 (2009).
55. W. Huang da, B. T. Sherman, R. A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* **4**, 44-57 (2009).
56. A. L. Bauernfeind *et al.*, Aerobic glycolysis in the primate brain: reconsidering the implications for growth and maintenance. *Brain Structure and Function* **219**, 1149-1167 (2014).
57. S. N. Vaishnavi *et al.*, Regional aerobic glycolysis in the human brain. *Proc Natl Acad Sci U S A* **107**, 17757-17762 (2010).
58. M. S. Goyal, M. Hawrylycz, J. A. Miller, A. Z. Snyder, M. E. Raichle, Aerobic glycolysis in the human brain is associated with development and neotenous gene expression. *Cell Metab* **19**, 49-57 (2014).
59. M. Li *et al.*, Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* **362**, eaat7615 (2018).
60. B. B. Biswal *et al.*, Toward discovery science of human brain function. *Proc Natl Acad Sci U S A* **107**, 4734-4739 (2010).
61. G. Cumming, *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis* (Routledge, 2013), pp. 207-230, 281-319.
62. T. Yarkoni, Big Correlations in Little Studies: Inflated fMRI Correlations Reflect Low Statistical Power—Commentary on Vul et al. (2009). *Perspectives on Psychological Science* **4**, 294-298 (2009).
63. M. D. Fox, M. Corbetta, A. Z. Snyder, J. L. Vincent, M. E. Raichle, Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proceedings of the National Academy of Sciences* **103**, 10046-10051 (2006).
64. M. Corbetta, G. L. Shulman, Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience* **3**, 201-215 (2002).

65. J. L. Vincent, I. Kahn, A. Z. Snyder, M. E. Raichle, R. L. Buckner, Evidence for a Frontoparietal Control System Revealed by Intrinsic Functional Connectivity. *Journal of Neurophysiology* **100**, 3328-3342 (2008).
66. S. Marek, N. U. F. Dosenbach, The frontoparietal network: function, electrophysiology, and importance of individual precision mapping. *Dialogues in clinical neuroscience* **20**, 133-140 (2018).
67. D. J. Felleman, D. C. Van Essen, Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* **1**, 1-47 (1991).
68. R. L. Buckner, J. R. Andrews-Hanna, D. L. Schacter, The brain's default network: anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences* **1124**, 1-38 (2008).
69. R. M. Braga, R. L. Buckner, Parallel Interdigitated Distributed Networks within the Individual Estimated by Intrinsic Functional Connectivity. *Neuron* **95**, 457-471 e455 (2017).
70. M. D. Fox *et al.*, The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 9673-9678 (2005).
71. M. Xia *et al.*, Reproducibility of functional brain alterations in major depressive disorder: Evidence from a multisite resting-state functional MRI study with 1,434 individuals. *Neuroimage* **189**, 700-714 (2019).
72. Y. Xie *et al.*, Alterations in Connectome Dynamics in Autism Spectrum Disorder: A Harmonized Mega- and Meta-Analysis Study Using the ABIDE Dataset. *bioRxiv* 10.1101/2021.10.03.462909, 2021.2010.2003.462909 (2021).
73. J. G. Nicholls *et al.*, *From neuron to brain, fifth edition* (Sinauer Associates Sunderland, MA, 2012), pp. 273-298.
74. X. Liang, Q. Zou, Y. He, Y. Yang, Coupling of functional connectivity and regional cerebral blood flow reveals a physiological basis for network hubs of the human brain. *Proc Natl Acad Sci U S A* **110**, 1929-1934 (2013).
75. Y. Chen, Q. Lin, X. Liao, C. Zhou, Y. He, Association of aerobic glycolysis with the structural connectome reveals a benefit-risk balancing mechanism in the human brain. *Proc Natl Acad Sci U S A* **118** (2021).
76. J.-P. Changeux, A. Danchin, Selective stabilisation of developing synapses as a mechanism for the specification of neuronal networks. *Nature* **264**, 705-712 (1976).
77. Z. Petanjek *et al.*, Extraordinary neoteny of synaptic spines in the human prefrontal cortex. *Proc Natl Acad Sci U S A* **108**, 13281-13286 (2011).
78. D. J. Miller *et al.*, Prolonged myelination in human neocortical evolution. *Proceedings of the National Academy of Sciences* **109**, 16480 (2012).
79. J. Hill *et al.*, Similar patterns of cortical expansion during human development and evolution. *Proc Natl Acad Sci U S A* **107**, 13135-13140 (2010).
80. C. Lebel *et al.*, Diffusion tensor imaging of white matter tract evolution over the lifespan. *Neuroimage* **60**, 340-352 (2012).
81. M. F. Paredes *et al.*, Extensive migration of young neurons into the infant human frontal lobe. *Science* **354** (2016).
82. K. Y. Kwan, N. Sestan, E. S. Anton, Transcriptional co-regulation of neuronal migration and laminar identity in the neocortex. *Development* **139**, 1535-1546 (2012).
83. Y. Zhu *et al.*, Spatiotemporal transcriptomic divergence across human and macaque brain development. *Science* **362** (2018).
84. I. Diez, J. Sepulcre, Neurogenetic profiles delineate large-scale connectivity dynamics of the human brain. *Nat Commun* **9**, 3876 (2018).
85. C. Cioli, H. Abdi, D. Beaton, Y. Burnod, S. Mesmoudi, Differences in Human Cortical Gene Expression Match the Temporal Properties of Large-Scale Functional Networks. *PLoS ONE* **9** (2014).

86. Y. Wei *et al.*, Statistical testing in transcriptomic-neuroimaging studies: A how-to and evaluation of methods assessing spatial and gene specificity. *Human Brain Mapping* https://doi.org/10.1002/hbm.25711 (2021).
87. M. Helmer *et al.*, On stability of Canonical Correlation Analysis and Partial Least Squares with application to brain-behavior associations. *bioRxiv* 10.1101/2020.08.25.265546 (2020).
88. R. L. Buckner, J. L. Roffman, J. W. Smoller, Brain Genomics Superstruct Project (GSP). Harvard Dataverse. https://doi.org/10.7910/DVN/25833.
89. J. B. Burt, M. Helmer, M. Shinn, A. Anticevic, J. D. Murray, Generative modeling of brain maps with spatial autocorrelation. *Neuroimage* **220**, 117038 (2020).
90. K. Brodmann, *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues* (Barth, 1909).
91. M. F. Glasser *et al.*, A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171-178 (2016).

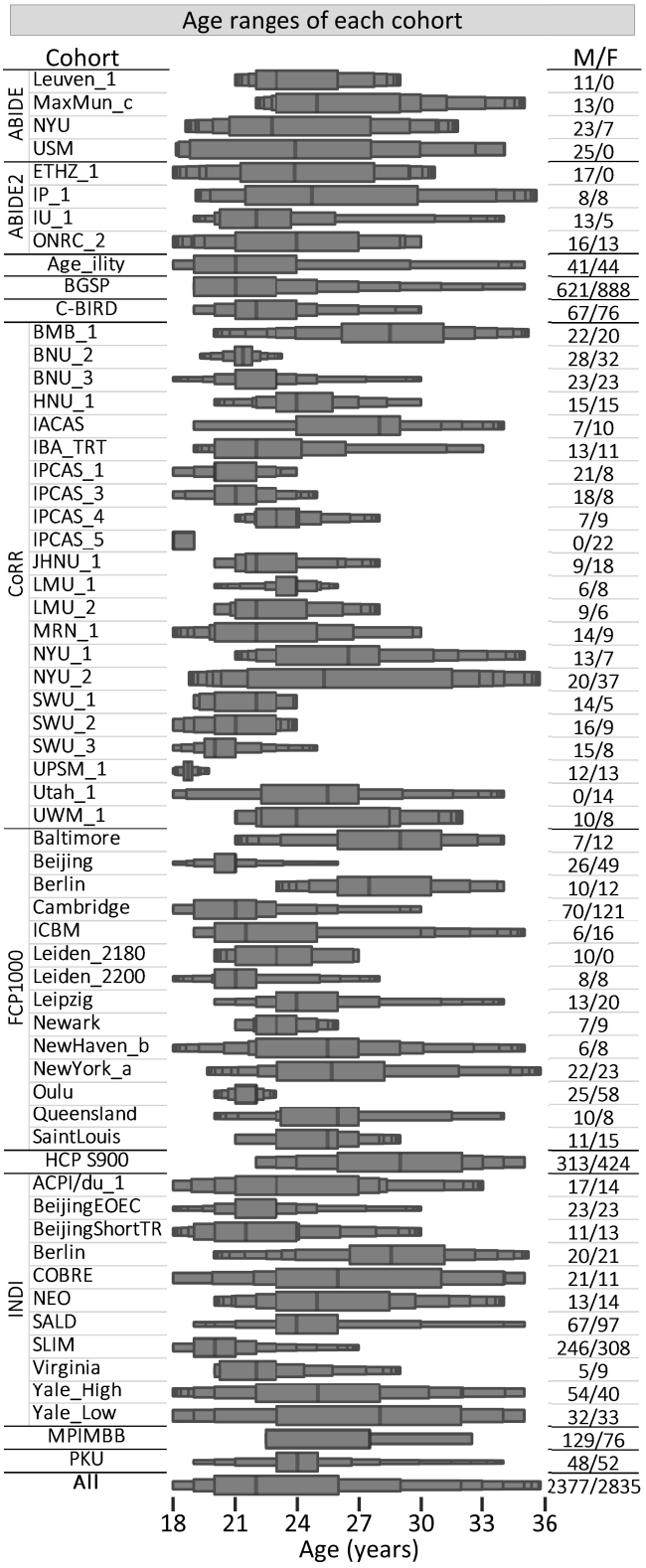


Figure 1. Enhanced box plot of the age ranges of each cohort. M/F, males/femals.

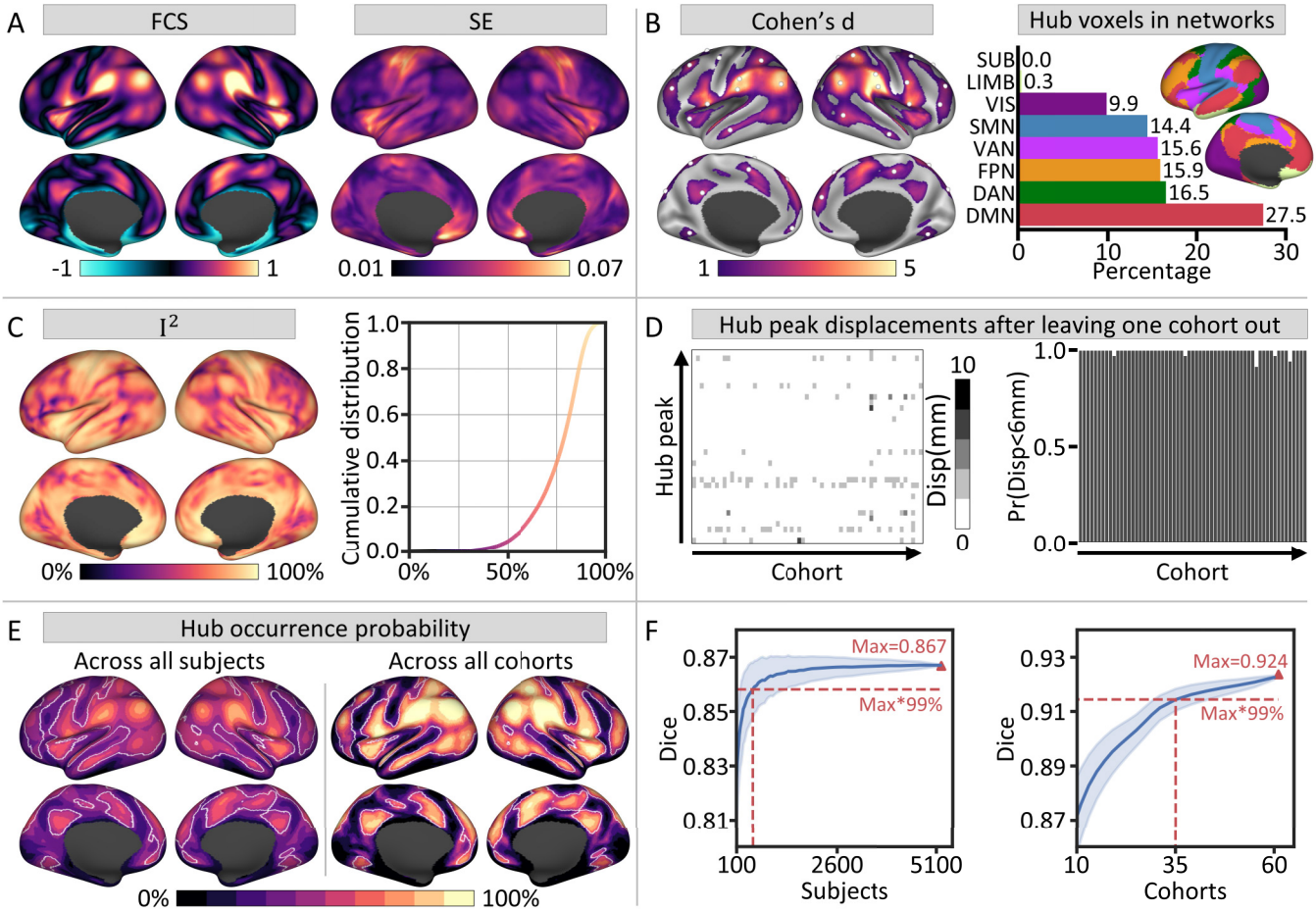
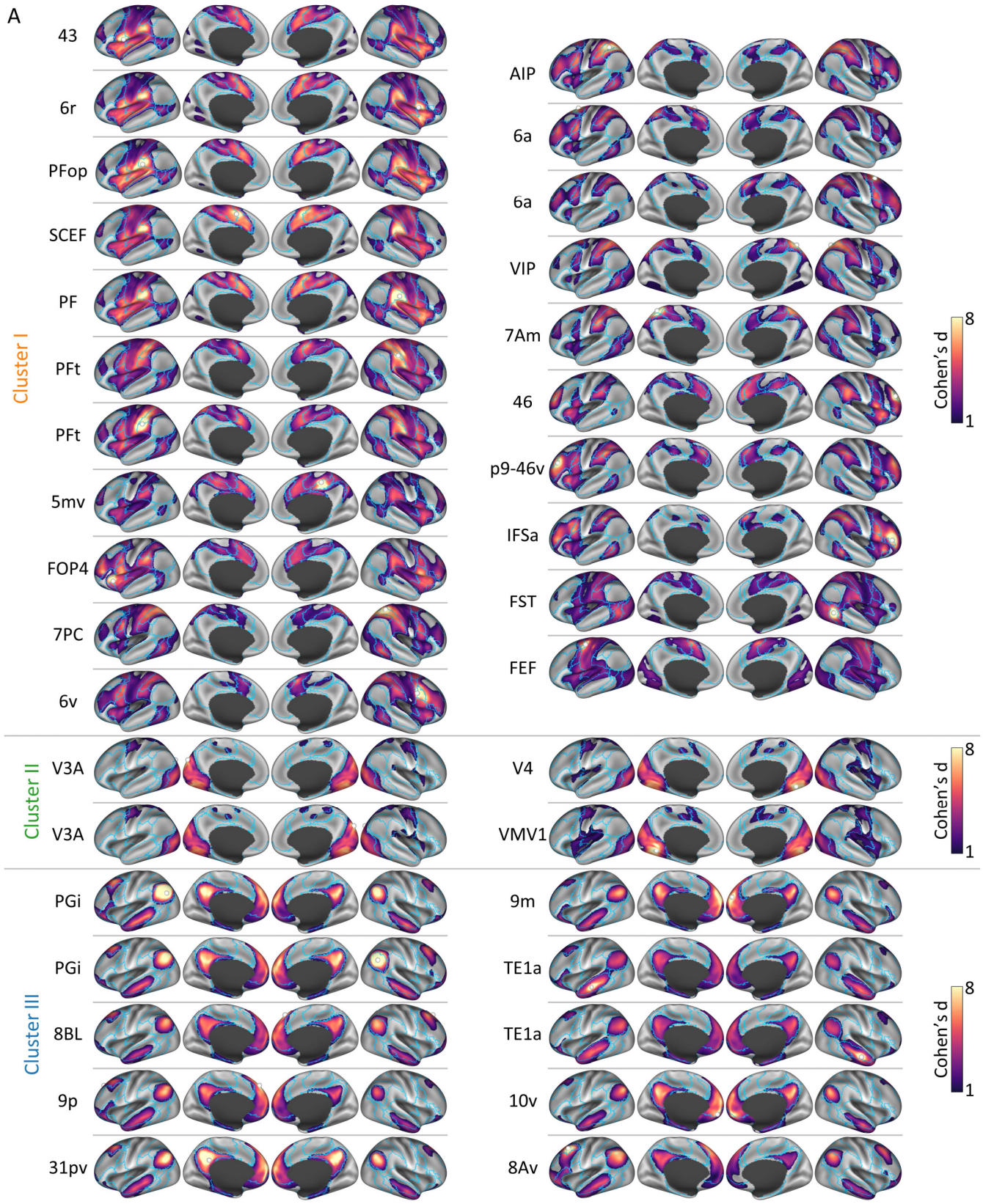


Figure 2. Highly consistent and reproducible functional connectome hubs. (A) Robust FCS pattern and its corresponding variance (standard error, SE) map estimated using a harmonized voxelwise random-effects meta-analysis across 61 cohorts. (B) Left: The most consistent functional connectome hubs ($p < 0.001$, 10,000 permutation tests, cluster size > 200 mm^3); white spheres represent hub peaks. Right: Hub voxel distribution in eight large-scale brain networks; insets, the seven large-scale cortical networks provided by (42) were rendered on the left hemisphere. (C) Left: Heterogeneity measurement I^2 estimated through the random-effects meta-analysis model. Right: Cumulative distribution function plot of I^2. (D) Left: Heatmap of displacement of the 35 hub peaks after leaving one cohort out. Right: Bar plot of the probability across the 35 hub peaks whose displacement was less than 6 mm after leaving one cohort out. (E) Hub occurrence probability map across all subjects (left) and all cohorts (right). White lines delineate boundaries of the identified hubs in Figure 2B. (F) Dice's coefficient of the identified hubs in Figure 2B compared with the top N (voxel number of the identified hubs in Figure 2B) voxels with the highest hub occurrence probability values across randomly selected subjects (left) and randomly selected cohorts (right). Blue shading represents the standard deviation across 2,000 random selectings.


A
Cluster I
43
6r
PFop
SCEF
PF
PFt
PFt
5mv
FOP4
7PC
6v
AIP
6a
6a
VIP
7Am
46
p9-46v
IFSa
FST
FEF
Cohen's d
8
1
Cluster II
V3A
V3A
V4
VMV1
Cohen's d
8
1
Cluster III
PGi
PGi
8BL
9p
31pv
9m
TE1a
TE1a
10v
8Av
Cohen's d
8
1

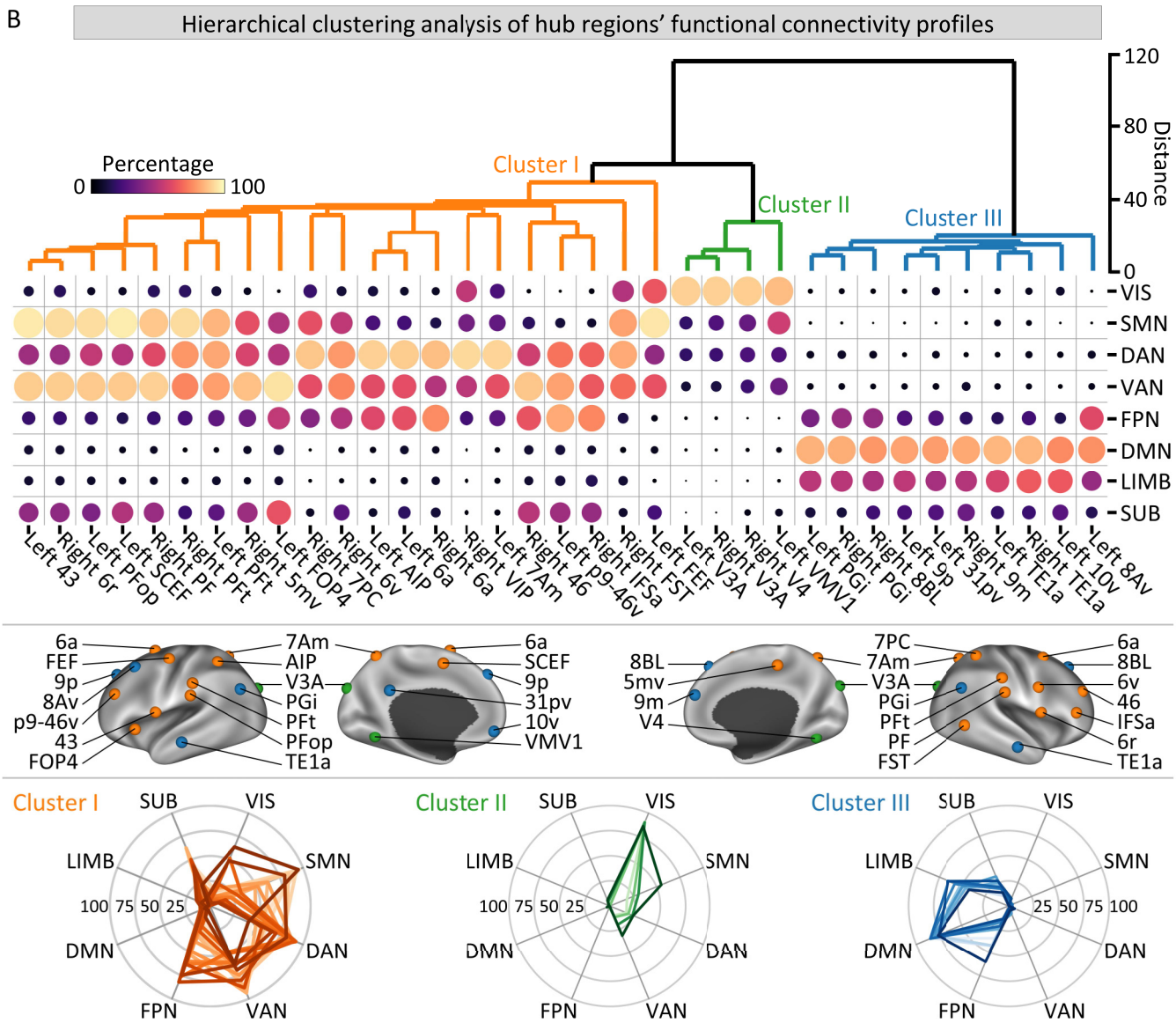


Figure 3. Connectivity profiles of connectome hubs. (A) Functional connectivity profiles of the 35 hubs. White spheres represent hub seeds. Blue lines delineate boundaries of the seven cortical networks shown in Figure 2B. (B) Top: Dendrogram derived by hierarchical clustering of the connectivity profile matrix. Middle: Thirty-five hubs were rendered using three different colors according to the hierarchical clustering solution. Bottom: Radar charts showing heterogeneous connectivity profiles of the three hub clusters.

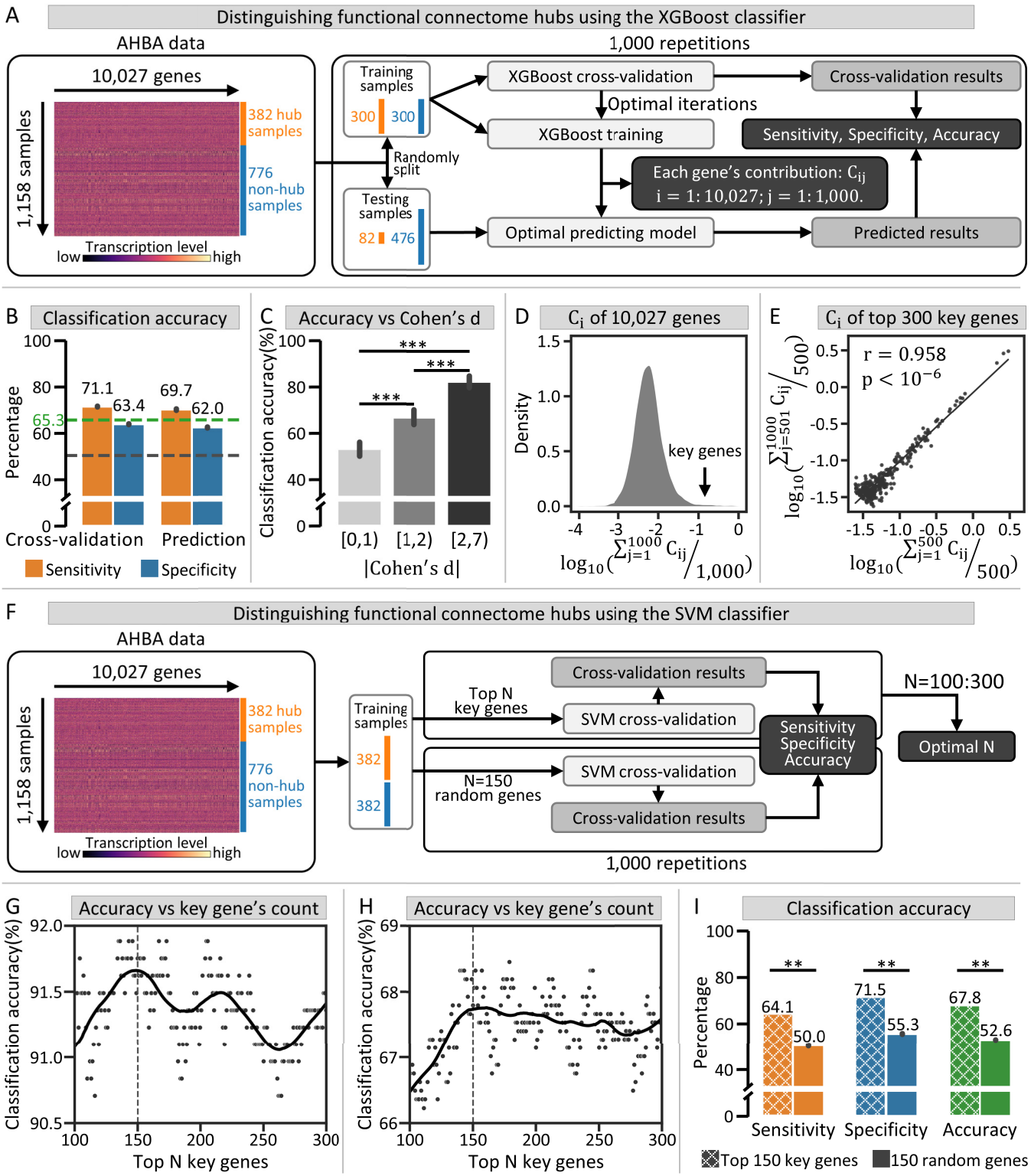


Figure 4. Transcriptomic data distinguish functional connectome hubs from non-hubs. (A) Schematic diagram of using the XGBoost model to classify brain samples as a hub or non-hub. (B) Accuracy rate of the XGBoost classifier. The horizontal gray dashed line represents the chance level accuracy rate (50%). The horizontal green dashed line represents the average

accuracy rate of the XGBoost classifier across 1,000 repetitions. (C) The rate to be correctly classified versus absolute Cohen's *d* value across 1,158 brain samples. (D) Density plot of 10,027 genes' logarithmic average contributions across 1,000 repetitions to the XGBoost classifier. Genes with the greatest contribution were considered key genes. (E) Regression plot of the logarithmic average contributions of the top 300 key genes across the first 500 repetitions versus those across the second 500 repetitions. Each dot represents one gene. (F) Schematic diagram of using the SVM model to classify brain samples as a hub or non-hub. (G and H) Accuracy rate of the SVM classifier versus the count of key genes used to distinguish 382 hub samples from 382 non-hub samples with the highest rate (G) or lowest rate (H) to be correctly classified by the XGBoost classifier. Each dot represents one SVM classifier. Black curves were estimated by locally weighted regression. (I) Bar plot showing the SVM classifier trained using top 150 key genes in (H) performing significantly better than those using 150 randomly selected genes. Gray bars in (B), (C), and (I) indicate 95% confidence intervals. Statistical significance in (C) was calculated by 10,000 permutation test with the spatial autocorrelations being corrected using a generative model (89). $^{**}p<0.001$; $^{***}p<0.0001$.

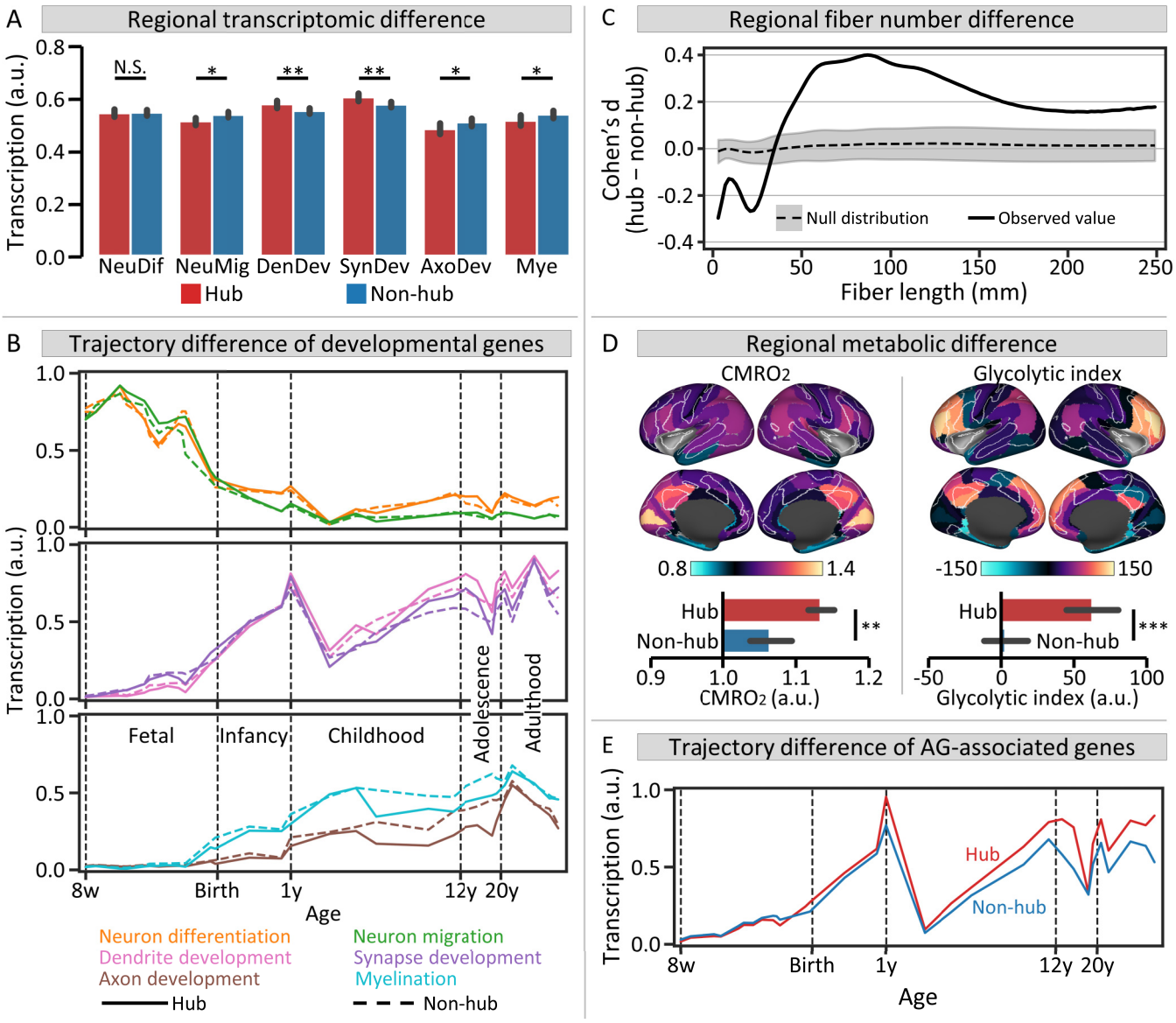


Figure 5. Spatiotemporal transcriptomic signatures of functional connectome hubs. (A) Regional transcriptomic differences between hub and non-hub samples for genes associated with key neurodevelopmental processes. We used the first principal component of each gene set's transcription level to plot and perform the statistical analysis. NeuDif, neuron differentiation; NeuMig, neuron migration; DenDev, dendrite development; SynDev, synapse development; AxoDev, axon development; Mye, myelination. (B) Transcriptomic trajectories of genes associated with key neurodevelopmental processes in hub regions (solid line) and non-hub regions (dashed line). (C) Fiber number difference between hub regions and non-hub regions across different fiber lengths. (D) Top left: Oxidative phosphorylation levels across Brodmann areas were measured by the $CMRO_2$. Bottom left: Bar plot of $CMRO_2$ in hub regions and non-hub regions. Top right: Aerobic glycolysis levels across Brodmann areas were measured by the glycolytic index. Bottom right: Bar plot of glycolytic index in hub regions and non-hub regions. White lines delineate boundaries of the identified hubs in Figure 2B. Data from the 82 Brodmann areas and seven subcortical regions were included in the alalysis. Brodmann areas with more

than 50% vertex or subcortical structures with more than 50% voxels identified as hubs were regarded as hub regions. (E) Transcriptomic trajectories of genes associated with aerobic glycolysis (AG) in hub regions (red line) and in non-hub regions (blue line). Transcriptomic trajectories in (B) and (E) were plotted using locally weighted regression by smoothing the first principal component of each gene set's transcription level against log2[postconception days]. For illustration purposes, we normalized the transcription level for each gene set across all brain samples to range from 0 to 1. Gray bars in (A) and (D) indicate 95% confidence intervals across brain samples. Statistical significance in (A) and (D) (*p<0.05; **p<0.005; ***p<0.0005; N.S.p>0.05) and the null distribution in (C) were calculated or constructed by 10,000 permutation tests with the spatial autocorrelations being corrected using a generative model (89). a.u., arbitrary unit; w, postconception week; y, postnatal year.

Table 1. Highly consistent functional connectome hubs using a harmonized meta-analysis.

No.	Hub	Location	MNI coordinates			Cohen's *d*	FCS	SE
			x	y	z			
1	Right PFt	PFt (superoanterior BA 40)	60	-21	45	6.267	1.072	0.022
2	Left PFt	PFt (superoanterior BA 40)	-60	-24	36	6.151	0.949	0.020
3	Right PF	PF (posterior BA 40)	60	-27	24	5.785	1.239	0.027
4	Left SCEF	Supplementary and cingulate eye field	0	0	51	5.635	1.000	0.023
5	Left PGi	PGi (inferior BA 39)	-51	-66	30	5.168	1.075	0.027
6	Left PFop	PF opercular (inferoanterior BA 40)	-63	-27	18	5.160	1.095	0.027
7	Left 43	Area 43	-57	3	3	4.927	1.114	0.029
8	Right 6r	Rostral area 6	57	6	0	4.916	1.184	0.031
9	Right PGi	PGi (inferior BA 39)	54	-60	30	4.739	1.007	0.027
10	Right 8BL	Area 8B lateral	21	36	51	4.655	0.713	0.020
11	Right 7PC	Area 7PC	36	-45	54	4.414	0.712	0.021
12	Left 9p	Area 9 posterior	-15	45	45	4.199	0.639	0.019
13	Right 6v	Ventral area 6	54	9	33	4.037	0.766	0.024
14	Left 8Av	Ventral area 8A	-39	18	48	3.990	0.561	0.018
15	Left AIP	Anterior intra-parietal area	-33	-45	45	3.474	0.567	0.021
16	Right FST	Fundus of the superior temporal area	54	-60	0	3.156	0.729	0.030
17	Right 9m	Area 9 middle	3	54	24	3.128	0.609	0.025
18	Left 31pv	Area 31p ventral	-3	-51	33	3.049	0.784	0.033
19	Right VIP	Ventral intra-parietal complex	18	-63	57	2.984	0.572	0.025
20	Right 6a	Area 6 anterior	33	3	63	2.975	0.454	0.020
21	Left FOP4	Frontal opercular area 4	-33	21	6	2.858	0.828	0.037
22	Right 5mv	Area 5m ventral	12	-30	45	2.822	0.701	0.032
23	Right 46	Area 46	36	42	30	2.779	0.656	0.030
24	Left 10v	Area 10v	0	57	-9	2.769	0.731	0.034
25	Left p9-46v	Area posterior 9-46v	-42	36	27	2.591	0.561	0.028
26	Left V3A	Area V3A	-15	-90	33	2.575	0.684	0.034
27	Left TE1a	Area TE1 anterior	-63	-15	-15	2.527	0.595	0.030
28	Right TE1a	Area TE1 anterior	60	-9	-21	2.494	0.580	0.030
29	Right IFSa	Anterior inferior frontal suleus	48	39	12	2.468	0.480	0.025
30	Left 7Am	Medial area 7A	-12	-60	60	2.461	0.475	0.025
31	Right V3A	Area V3A	18	-87	36	2.442	0.645	0.034
32	Right V4	Fourth visual area	24	-63	-9	2.339	0.446	0.024
33	Left 6a	Area 6 anterior	-24	3	63	2.317	0.331	0.018
34	Left VMV1	Ventromedial visual area 1	-18	-60	-6	1.937	0.397	0.026
35	Left FEF	Frontal eye fields	-45	-9	57	1.412	0.640	0.058

Note: Hub name and location description were based on (90, 91). BA: Brodmann area.