

RESEARCH ARTICLE

Open Access



Evolution and transition of expression trajectory during human brain development

Ming-Li Li^{1,2†}, Hui Tang^{3†}, Yong Shao^{1,2}, Ming-Shan Wang^{1,2}, Hai-Bo Xu^{1,2}, Sheng Wang¹, David M. Irwin^{1,4,5}, Adeniyi C. Adeola^{1,2}, Tao Zeng³, Luonan Chen^{3,6,7*†} , Yan Li^{8*†} and Dong-Dong Wu^{1,2,7*†} 

Abstract

Background: The remarkable abilities of the human brain are distinctive features that set us apart from other animals. However, our understanding of how the brain has changed in the human lineage remains incomplete, but is essential for understanding cognition, behavior, and brain disorders in humans. Here, we compared the expression trajectory in brain development between humans and rhesus macaques (*Macaca mulatta*) to explore their divergent transcriptome profiles.

Results: Results showed that brain development could be divided into two stages, with a demarcation date in a range between 25 and 26 postconception weeks (PCW) for humans and 17-23PCW for rhesus macaques, rather than birth time that have been widely used as a uniform demarcation time of neurodevelopment across species. Dynamic network biomarker (DNB) analysis revealed that the two demarcation dates were transition phases during brain development, after which the brain transcriptome profiles underwent critical transitions characterized by highly fluctuating DNB molecules. We also found that changes between early and later brain developmental stages (as defined by the demarcation points) were substantially greater in the human brain than in the macaque brain. To explore the molecular mechanism underlying prolonged timing during early human brain development, we carried out expression heterochrony tests. Results demonstrated that compared to macaques, more heterochronic genes exhibited neoteny during early human brain development, consistent with the delayed demarcation time in the human lineage, and proving that neoteny in human brain development could be traced to the prenatal period. We further constructed transcriptional networks to explore the profile of early human brain development and identified the hub gene *RBFOX1* as playing an important role in regulating

(Continued on next page)

* Correspondence: lnchen@sibs.ac.cn; liyan0910@ynu.edu.cn; wudongdong@mail.kiz.ac.cn

†Ming-Li Li, Hui Tang, Luonan Chen, Yan Li and Dong-Dong Wu contributed equally to this work.

³State Key Laboratory of Cell Biology, Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Shanghai 200031, China

⁸State Key Laboratory for Conservation and Utilization of Bio-Resource, Yunnan University, Kunming 650091, Yunnan, China

¹State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, Yunnan, China

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(Continued from previous page)

early brain development. We also found *RBFOX1* evolved rapidly in its non-coding regions, indicating that this gene played an important role in human brain evolution. Our findings provide evidence that *RBFOX1* is a likely key hub gene in early human brain development and evolution.

Conclusions: By comparing gene expression profiles between humans and macaques, we found divergent expression trajectories between the two species, which deepens our understanding of the evolution of the human brain.

Keywords: Humans, Macaques, Expression trajectory, Transcriptome, Brain evolution

Background

Our highly developed and distinctive brains, which set humans apart from other mammals, are the product of evolution [1, 2], the mechanism of which has fascinated people for centuries [3]. Based on compelling differences in cognitive and behavioral capacities, but relatively close phylogenetic relationship between humans and non-human primates (NHPs) [4–6], recent comparative analyses have provided a novel strategy to study human-specific neurodevelopment [7–9]. Increasingly persuasive evidence suggests that brain development is not static but is a continuous process of molecular changes throughout life, including changes in gene expression, glucose metabolism, and synaptic density [10–14]. Previous comparative analyses between humans and NHPs have only offered a snapshot in time [15, 16]. However, it is necessary to compare the whole process of brain development to provide a more objective and comprehensive understanding of human brain evolution.

Earlier research noted that neurodevelopmental timing is impacted by different developmental rates and life history strategies [2]. These differences in neurodevelopmental timing among species, also called heterochrony, have long been considered a crucial impetus for evolution [17–19]. Humans have an unusually extended childhood and slow rate of neurodevelopment (known as neoteny) relative to other animals, which is considered a possible mechanism for human brain evolution [18, 20]. While previous studies have primarily focused on heterochronic gene expression during postnatal brain development [20], the macroscopic layout of the brain is nearly complete at the time of birth [21]. Thus, extending comparative analysis to the prenatal stages is necessary for exploring the features of neurodevelopment.

Currently, it is widely accepted that changes in spatio-temporal gene expression play a critical role in the emergence of the sophisticated human brain, and several attempts have been made to estimate divergent gene expression patterns between humans and NHPs [15, 22, 23]. However, our understanding of how gene expression patterns have changed in the human lineage remains incomplete. With increasing high-quality brain transcriptome data [24–26], an unprecedented opportunity to

investigate gene expression trajectory in multiple brain regions and different developmental stages among primates has become possible [22, 27]. In this study, we collected large-scale gene expression data from humans and macaques to systematically investigate and compare their divergent gene expression trajectory. We aimed to identify critical states during brain development as well as novel molecular mechanisms underlying human brain evolution.

Results

Study design

Figure 1 highlights the strategy used to investigate evolution of gene expression trajectory in humans, including hierarchical clustering and dynamic network analyses to identify demarcation times of brain development in humans and macaques, expression heterochrony analysis to explore the mechanism of neurodevelopmental timing in humans, and differential expression and gene co-expression network analyses to identify key genes in human brain development and evolution.

Human brain transcriptome data were used in this study, including RNA-sequencing (RNA-seq) and microarray data across multiple brain regions downloaded from the Allen Brain Atlas [24, 25] (Table 1; Additional file 1: Table S1 ~ Table S2). These data cover 14 brain regions spanning 27 different developmental ages from 8 postconception weeks (PCW) to postnatal 40 years old (Table 2). In total, 30,881 and 17,280 genes had detectable expression signals in the RNA-seq and microarray data, respectively. We also used microarray data from macaques, which contained five brain regions (22 brain subregions) corresponding to brain regions in humans (Table 1; Additional file 1: Table S3 ~ Table S4) and spanning 8 PCW to postnatal 48 months (Table 2). In total, 15,381 genes exhibited detectable expression signals.

Different developmental trajectories and demarcation times in humans and macaques

We performed hierarchical clustering analysis based on gene expression levels to determine whether transformation exists during brain development. In humans, the RNA-seq and microarray data supported the division of

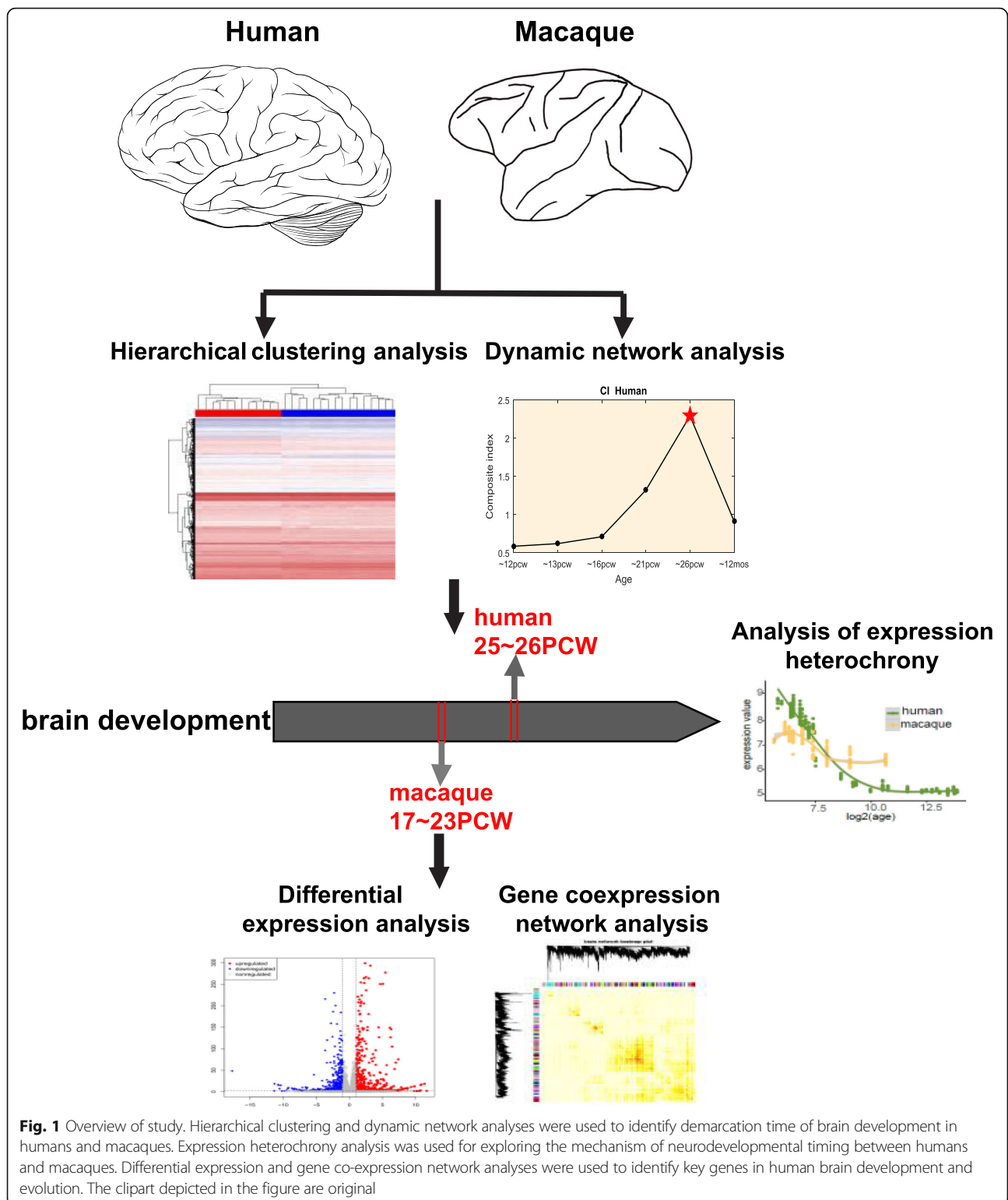


Table 1 Human and macaque tissues used in this study

Human	A1C	Primary auditory cortex	
	AMY	Amygdaloid complex	
	MD	Mediodorsal nucleus of thalamus	
	DFC	Dorsolateral prefrontal cortex	
	HIP	Hippocampus	
	IPC	Posteroventral (inferior) parietal cortex	
	ITC	Inferolateral temporal cortex	
	M1C	Primary motor cortex	
	ACC	Anterior (rostral) cingulate cortex	
	OFC	Orbital frontal cortex	
	STC	Posterior (caudal) superior temporal cortex	
	STR	Striatum	
	V1C	Primary visual cortex	
	VFC	Ventrolateral prefrontal cortex	
	Macaque	HIP	CA1or (stratum oriens of CA1) CA1py (stratum pyramidale of CA1) CA1ra (stratum radiatum of CA1) CA2py (stratum pyramidale of CA2) CA3py (stratum pyramidale of CA3) DGgr (granular layer of dentate gyrus) DGpf (polyform layer of dentate gyrus) DGsg (subgranular zone of dentate gyrus)
		STR	NAC (nucleus accumbens) ic (internal capsule) Pu (putamen) (Gpe) External segment of globus pallidus Internal segment of globus pallidus
ACC		(rCG2) Layer II of rostral cingulate cortex (rCG3) Layer III of rostral cingulate cortex (rCG5) Layer V of rostral cingulate cortex (rCG6) Layer VI of rostral cingulate cortex	
V1C		(V1–3) Layer III of V1 (V1–4) Layer IVA of V1 (V1–5) Layer V of V1	
AMY		(Me) Medial nucleus of amygdaloid (PL) Paralaminar nucleus	

The brain regions marked with bold in human represent that coexist with macaque

brain development into two stages, with a demarcation time of 25–26 PCW (Fig. 2a; Additional file 2: Figure S1). In macaques, gene expression levels from the microarray data in most brain regions were also clustered into two groups, with a demarcation time of 17 PCW ~ birth (17–23PCW)(Fig. 2a; Additional file 2: Figure S2). These results suggest that brain development in both humans and

macaques could be divided into two major stages, separated by species-specific demarcation points that occurred prior to birth (25–26 PCW in humans and 17–23 PCW in macaques), rather than birth times, which have been widely used as a uniform demarcation time of neurodevelopment across species [28, 29].

Transitional state and critical transitions during brain development based on dynamic network biomarker (DNB) analysis

Having identified the demarcation points of brain development in humans and macaques, we further applied DNB analysis to verify whether the above demarcation points were also transitional states of brain development. Based on nonlinear dynamic theory, biological processes, such as brain development, are not always smooth, but can exhibit dramatic transitions from one state to another [30, 31]. When the process occurs near a critical transition phase, a dominant group of genes/molecules, i.e., DNBs, can drive transition of the dynamic process [30, 32, 33]. We thus performed a genome-wide DNB analysis to identify the transitional states and DNB genes in both humans and macaques.

The DNB results demonstrated that the transitional states of brain development in humans and macaques occurred at around 26 PCW and around 17PCW, respectively (Fig. 2b-c; Additional file 2: Figure S3). After the transitional state, gene expression patterns changed markedly. The transitional states identified by DNB analysis were largely consistent with the demarcation points above (Fig. 2a), thus supporting the robustness of our results. We also obtained the corresponding DNBs, which included 369 DNB genes in humans and 34 DNB genes in macaques (Additional file 1: Table S5). The greater number of DNB genes in humans suggests more dramatic changes between the early and later stages of brain development in humans relative to that in macaques.

Transcriptional profile and cell fate change from early to later stages of brain development

We used differential expression analysis to compare the degree of change in early and later gene expression between humans and macaques. We selected five brain regions (i.e., hippocampus (HIP), striatum (STR), anterior cingulate cortex (ACC), amygdala (AMY), and primary visual cortex (V1C)) that coexist and contain similar sample sizes in the two species. Based on microarray probes, which matched between humans and macaques, we found a larger number of differentially expressed genes (DEGs) (Benjamini-Hochberg FDR < 0.05, fold change [FC] > 1.5) between early and later stages in humans than in macaques (Fig. 3a; Additional file 1: Table S6). These results also suggest more dramatic

Table 2 Age distribution of humans and macaques

Human		Macaque	
Early period	Later period	Early period	Later period
56 PCD (8 PCW)	182 PCD (26 PCW)	50 PCD (8 PCW)	0 m
72 PCD (9 PCW)	4 m	70 PCD (10 PCW)	3 m
84 PCD (12 PCW)	10 m	80 PCD (12 PCW)	12 m
91 PCD (13 PCW)	1 y	90 PCD (13 PCW)	48 m
112 PCD (16 PCW)	2 y	120 PCD (17 PCW)	
119 PCD (17 PCW)	3 y		
133 PCD (19 PCW)	4 y		
147 PCD (21 PCW)	8 y		
168 PCD (24 PCW)	13 y		
175 PCD (25 PCW)	15 y		
	18 y		
	21 y		
	23 y		
	30 y		
	36 y		
	37 y		
	40 y		

PCW postconception week, PCD postconception days, m month, y year

changes between early and later stages of brain development in humans relative to macaques.

To better understand brain developmental processes in the human lineage, we conducted functional enrichment analysis of the DEGs between early and later stages in humans. Results showed that up-regulated genes in the early stage were mainly involved in cell cycle, DNA packaging, and meiosis (Fig. 3b), whereas, up-regulated genes in the later stage were enriched in synaptic signaling, myelination, and axon establishment (Fig. 3c). Remarkably, the DEG patterns well matched the reported properties of the neurodevelopmental timeline in humans [35].

Increasing evidence suggests that a cell fate switch from neurons to glial cells is operational in prenatal brains and represents a key process in brain development [36–38]. We thus considered whether the demarcation points in human correspond to the transient time of known neuron to glial cell fate switch. We found up-regulated genes in the early stage were predominantly enriched in neuronal genes (Fig. 3d), whereas up-regulated genes in the later stage represented a diversity of cell-type-associated genes, including astrocytes,

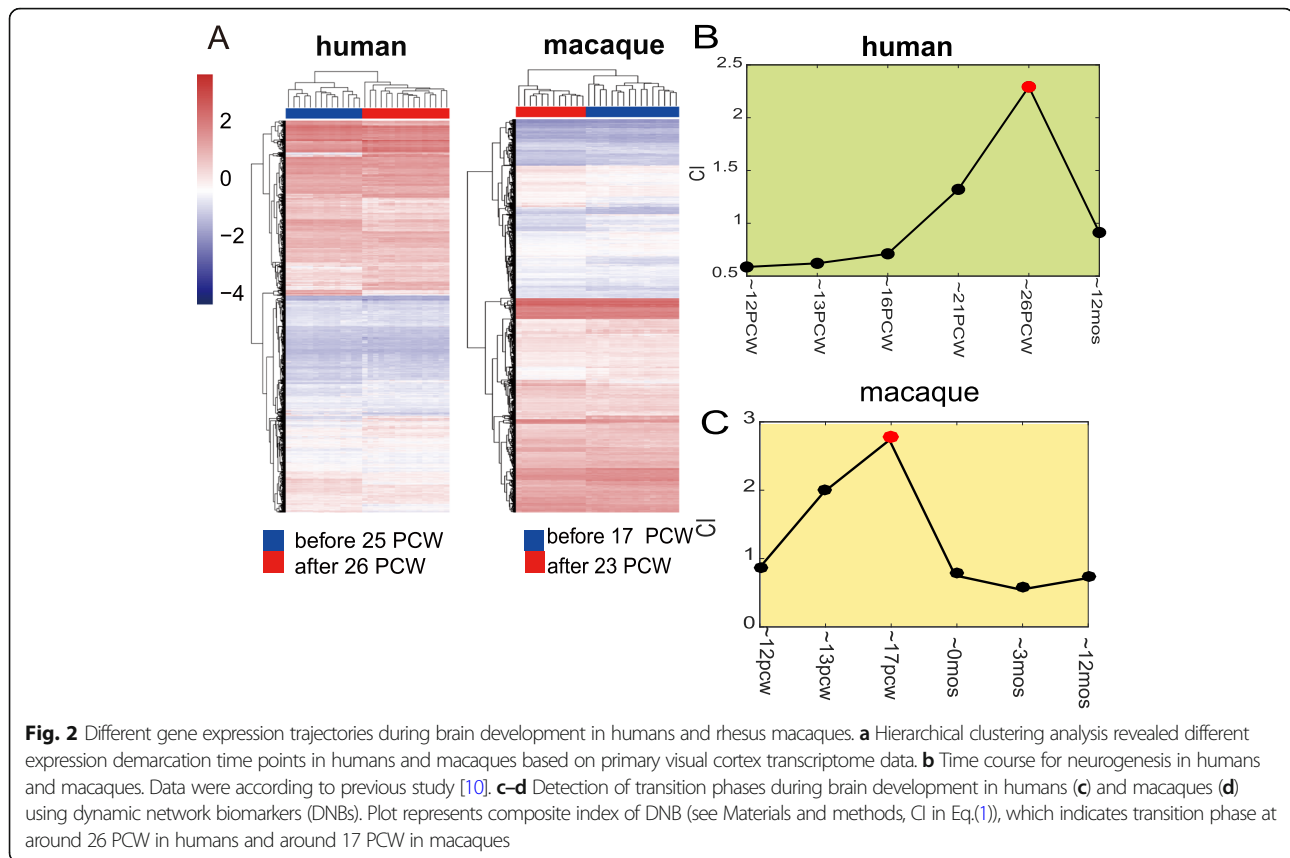
oligodendrocytes, and neurons (Fig. 3d), which likely reflects the transformation of cell fate switch from neurons to glial cells during the demarcation point. This conclusion is confirmed by the previous single-cell transcription analysis [39], which reported that neurons developed from neural progenitor cells in early gestational weeks (GW8, GW9, GW10, GW12, GW13, GW16, GW19, GW23), whereas glial cells (oligodendrocyte progenitor cells and astrocytes) differentiated from neural progenitor cells in later weeks (GW26).

Earlier studies have also reported several pathways that govern the neuron to astrocyte cell switch, including the gp130/JAK/STAT and MEK/MAPK pathways ... [38, 40]. We found that DEGs between early and later stages were significantly enriched in the MEK/MAPK pathway ($P = 4.6e-04$, Fisher's exact test) (Additional file 1: Table S7). Although DNBs were not significantly enriched in the MEK/MAPK pathway ($P = 0.15$, Fisher's exact test), 15 DNBs were still involved (Additional file 1: Table S7). This suggests that the MEK/MAPK pathway likely plays an important role in the transformation ratio of cell type from neurons to glial cells during the human demarcation point (25–26 PCW).

Molecular mechanism underlying protracted timing of early human brain development

The different demarcation points (25–26 PCW in humans and 17–23PCW in macaques) identified here reflect prolonged timing of early brain development in the human lineage. Thus, we performed heterochronic analysis to explore the molecular mechanism underlying different timing of early neurodevelopment between humans and macaques.

We again used five brain regions (i.e., HIP, STR, ACC, AMY, and VIC) that coexist in humans and macaques to test heterochronic gene expression. After rigorous quality control (see Methods), we retained 9758 genes with microarray probes well matched between humans and macaques. Among these genes, we selected several that showed both age-related and species-specific differences for each brain region (see Methods; Additional file 1: Table S8). We then sorted these genes into two categories: (i) human acceleration genes, whose expression change was significantly faster during human brain development than that during macaque brain development (Fig. 4a); and, (ii) human neoteny genes, whose expression change was significantly delayed during human brain development than that during macaque brain development (Fig. 4b), as defined in previous study [20]. Compared to macaques, more genes displayed a neotenic pattern in all five human brain regions (Fig. 4c; Additional file 1: Table 9), consistent with the above delayed demarcation point in the human lineage. In



addition, the result suggest that neoteny of human brain development could be traced to prenatal period.

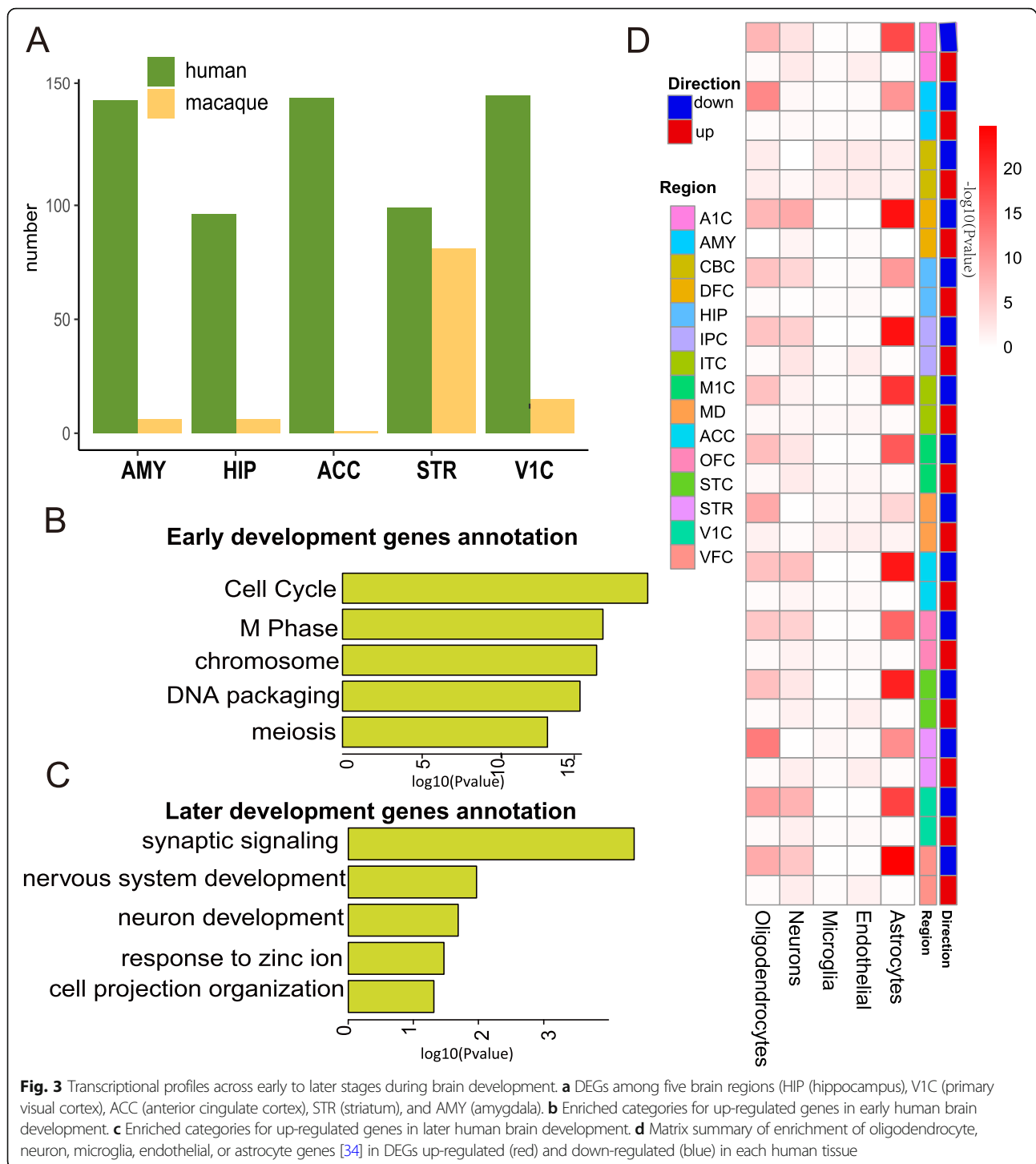
Interestingly, the neotenic genes from the five brain regions significantly overlapped (Additional file 1: Table 10), suggesting that neotenic mechanisms among different brain regions are largely convergent. The functions of these neotenic genes were mainly involved in neurodevelopment-related pathways (Fig. 4d), suggesting that more neurodevelopmental genes exhibited neotenic features, which eventually help humans develop a more complex brain.

Co-expression analysis identifies gene network during early brain development in humans

Extended timing of early neurodevelopment in humans is important for brain evolution [2, 18]. We applied weighted gene co-expression network analysis (WGCNA) to further explore the transcriptional profile of early neurodevelopment in humans [41–43]. A total of 38 modules related to early human neurodevelopment were identified (see Methods; Fig. 5a; Additional file 1: Table S11). To quantify network reorganization across early and later brain development, we applied modular differential connectivity

(MDC), which is the ratio of the average connectivity for any pair of modules sharing genes in the early stage compared to that in the later stage. Among the 38 early modules, five (GCM1–GCM5) showed gain of connectivity compared to later development, with co-regulation enhanced between genes in these modules. In contrast, 21 modules (LCM1–LCM21) showed loss of connectivity and 12 modules (NCM1–NCM12) (31.5%) showed no change in connectivity and were conserved during development (Additional file 2: Figure S4A–S4B).

For modules with gains or losses of connectivity, we ranked them according to their degree of DEG enrichment across brain regions (Fig. 5a) and MDC scores (Additional file 1: Table S11). Module GCM1, which showed a gain of connectivity in the early stage, was identified as the most highly ranked module. The genes in this module were enriched in neurogenesis, neuron projection morphogenesis, and axon development (Fig. 5d). Additionally, GCM1 showed highly significant enrichment for known autism susceptibility markers ($P = 5.49E-07$; Fisher's exact test) [46], and the expression levels of genes in this module were significantly higher during early brain development (Additional file 2: Figure S4C). These results suggest that GCM1 likely

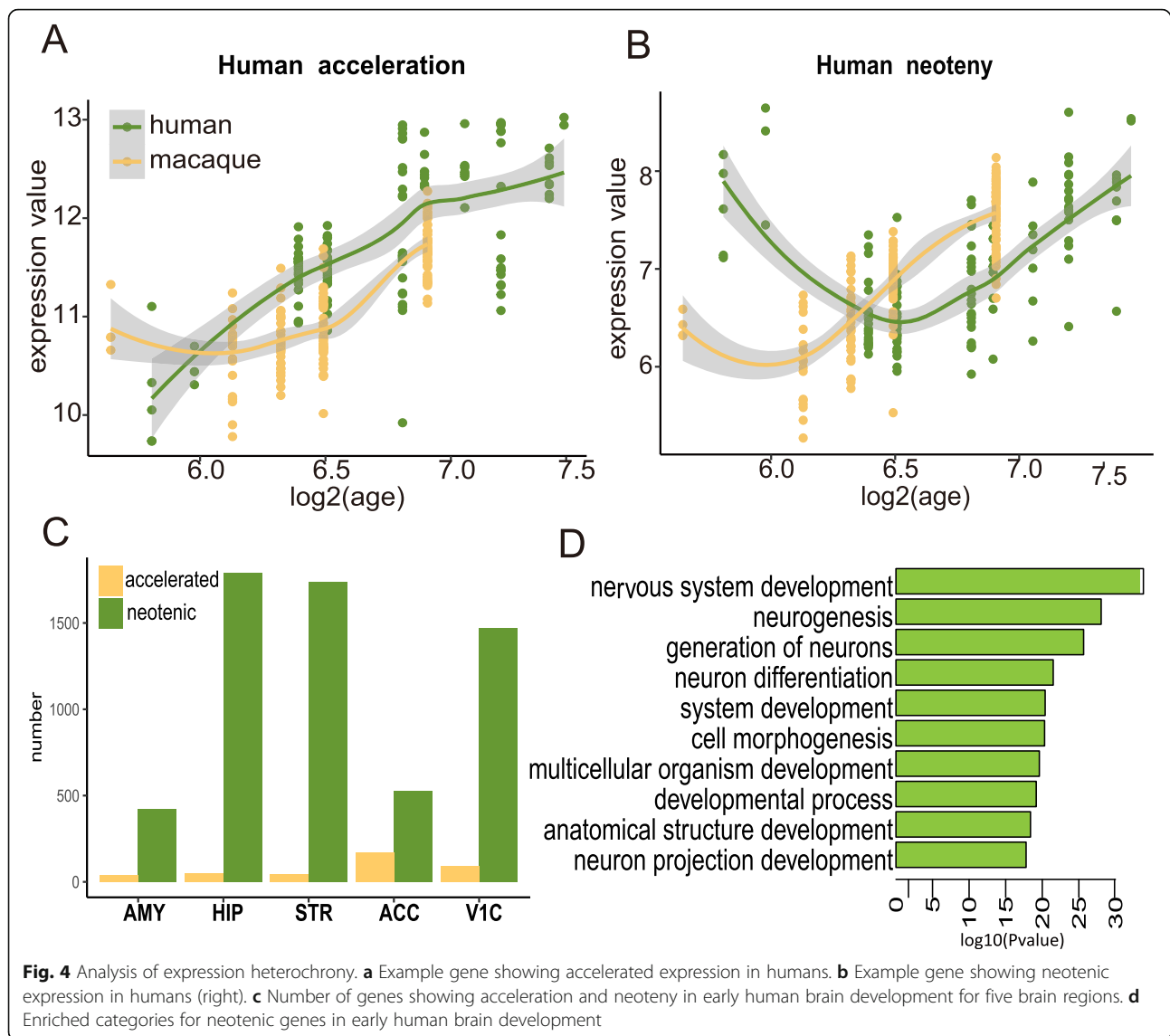


plays an important role in early development of the human brain.

Remarkably, most genes in the GCM1 module were located in human-accelerated conserved non-coding sequences (HACNSs) ($P = 3.19E-16$; Fisher's exact test) [44] or in human DNA sequence accelerated regions

(HARs) ($P = 7.49E-12$; Fisher's exact test) [45] (Fig. 5c), suggesting that genes in the GCM1 module also likely played an important role in human brain evolution.

We next mapped the genes in GCM1 to single-cell expression data derived from 20,262 prenatal human prefrontal cortex cells that ranged in age from 8 to 26



gestational weeks [39] (Additional file 2: Figure S5), which represent a broad diversity of cell types including neural progenitor cells, interneurons, astrocytes, oligodendrocyte progenitor cell, microglia and excitatory neurons. The expression patterns of the GCM1 genes closely matched the cell type specific of excitatory neurons (see Methods) (Fig. 5e), confirming that genes in the GCM1 module function through excitatory neurons.

We further reconstructed the network structure of genes within the GCM1 module based on their connectivity and identified 53 hub genes, 36 of which were early stage-specific hub genes [47] (see Methods, Fig. 6a). Among these genes, *RBFOX1* (RNA Binding Fox-1 Homolog 1) was of particular interest. *RBFOX1* is a highly conserved splicing regulator that displays higher expression in the brain than in other tissues [48] (Fig. 6b). *RBFOX1* is implicated in autism, epilepsy

syndromes, and Alzheimer's disease [49–51] and plays an important role in mammalian brain development [52]. Interestingly, *RBFOX1* is also a DNB gene, and therefore considered to play an important role in critical transition during brain development [30].

We next used evolutionary analysis to test if *RBFOX1* experienced positive selection in the human lineage. Although the protein coding sequence of *RBFOX1* has not changed in humans compared to other primates (Additional file 2: Table S12; see Methods), six HARs were found in the non-coding regions for this gene (Fig. 6c). HARs are non-coding regions conserved across mammals, and which have acquired many sequence changes in humans since their divergence from chimpanzees [45]. Only seven human RefSeq genes from the entire human genome (based on hg18) contain six or more HARs [53], suggesting that strongly human-specific accelerated

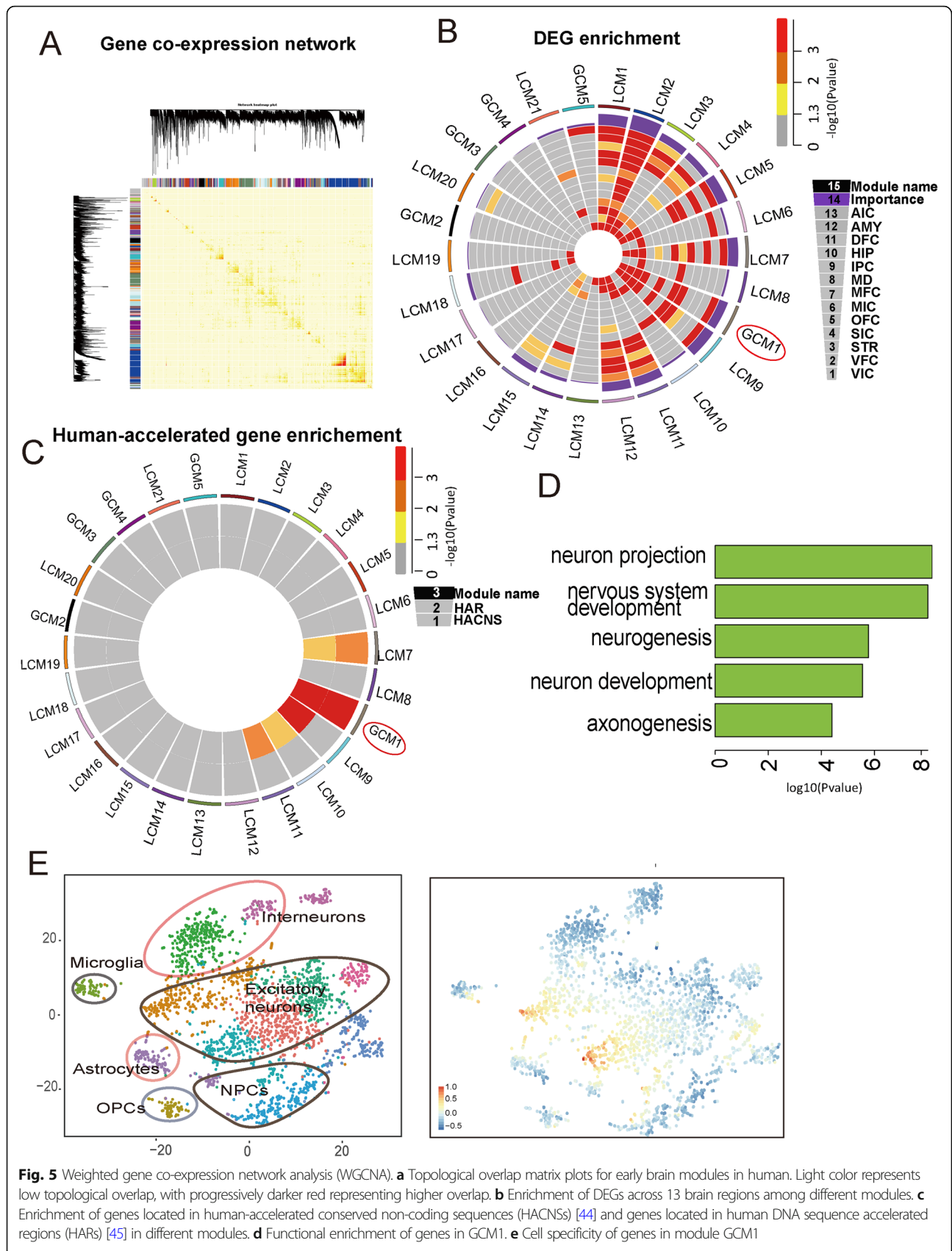


Fig. 5 Weighted gene co-expression network analysis (WGCNA). **a** Topological overlap matrix plots for early brain modules in human. Light color represents low topological overlap, with progressively darker red representing higher overlap. **b** Enrichment of DEGs across 13 brain regions among different modules. **c** Enrichment of genes located in human-accelerated conserved non-coding sequences (HACNSs) [44] and genes located in human DNA sequence accelerated regions (HARs) [45] in different modules. **d** Functional enrichment of genes in GCM1. **e** Cell specificity of genes in module GCM1

evolution occurred recently in the non-coding region of the human *RBFOX1* gene. Our findings provide evidence that *RBFOX1* is a likely key hub gene in early human brain development and evolution. In addition, *RBFOX1* also showed cell specificity to excitatory neurons by single cell transcriptome analysis (Fig. 6d), suggesting that *RBFOX1* functions through excitatory neurons.

Discussion

The remarkable abilities of the human brain set us apart from NHPs. With the advent of large-scale genomic, transcriptomic, and epigenomic data, many genetic underpinnings of the rapid evolution of the human brain have been revealed [54–58]. However, our understanding of how the brain has changed in the human lineage remains incomplete [3]. Based on large-scale transcriptomic and genomic data, the results of the current study provide new insight into the evolution and transition of gene expression trajectory in the human brain.

Firstly, we found that brain development could be divided into two stages in both humans and macaques; more specifically, demarcation times of 25–26 PCW and 17–23 PCW in humans and rhesus macaques, respectively. Further DNB analysis indicated that the demarcation points were nearly the same as the critical transitional states during brain development in humans and macaques. Previous studies on brain development have primarily used birth as the default boundary [28, 29, 59]. However, we suggest that the demarcation points identified here should be considered in the future to minimize biases in studies of brain development.

Secondly, we also found that neoteny of human brain development could be traced to the prenatal period. Previous studies have primarily focused on heterochronic gene expression during postnatal brain development [7]. The macroscopic layout of the brain is nearly complete at the time of birth [8]. Thus, extending comparative analysis to the prenatal stages is necessary to explore the features of neurodevelopment, which is lacking in prior studies. Thus, in this paper, we performed heterochronic analysis across prenatal samples between humans and macaques, and found that more genes displayed a neotenic pattern in humans than in macaques, consistent with the delayed demarcation time in the human lineage, and proving that neoteny in human brain development could be traced to the prenatal period.

Thirdly, we used gene co-expression network analysis to identify transcription profiles in early human brain development and identified that the *RBFOX1* gene likely plays an important role in early human brain development and displays positive selection in its non-coding region [50–52]. Therefore, we speculated that *RBFOX1* is a likely key hub gene in early human early brain

development and evolution. As such, we propose that *RBFOX1* should be considered in further neurodevelopmental research.

Finally, we highlighted the importance of excitatory neurons in human brain development and evolution. Over the past few decades, the comparisons of excitatory neurons between humans and NHPs have mainly focused on their differences in morphology and abundance [60, 61]. Further molecular biology research on excitatory neurons is limited. In this paper, we found that the GCM1 module and *RBFOX1* gene were related to early human brain development and evolution and were enriched in excitatory neurons. Therefore, studies on excitatory neurons would be promising for exploring human brain evolution.

We note that the study presented here is far from being comprehensive: Firstly, Based on currently available transcriptome data, we identified a demarcation line time-frame of brain development in humans and macaques. The precise demarcation point could not be concluded from existing data but should be explored in future studies.

Secondly, due to the relatively small sample sizes used in the current study, as well as the sparse distribution of samples across ages, we cannot rule out certain important changes in transcriptional profiles during neurodevelopment that may have occurred beyond the sampling range used in this study. For instance, previous studies have reported that during juvenile development in humans (1–8 postnatal years), the cerebral cortex consumes nearly twice the amount of glucose as observed during adulthood and is accompanied by dramatic changes in synaptic density during that developmental window [11, 12]. Thus, the transition state we identified is not absolute, with more saturated samples across different ages required to confirm our conclusions.

Thirdly, comparative analysis between humans and macaques was based on microarray data only, which rely on pre-existing knowledge of RNA sequences; as such, some important genes may be missed.

Finally, due to the lack of prenatal transcriptome data on brain development in hominoids, it is difficult to compare hominoids with humans, which would be valuable when exploring human brain evolution.

Further analyses, including expression data analysis across additional development stages, comparative analysis of RNA-seq data between humans and NHPs, as well as analysis incorporating more hominoids, are needed to expand our results.

Conclusions

In this study, we integrated transcriptomic analysis to reveal the evolution and transition of expression trajectories during human brain development. By comparing

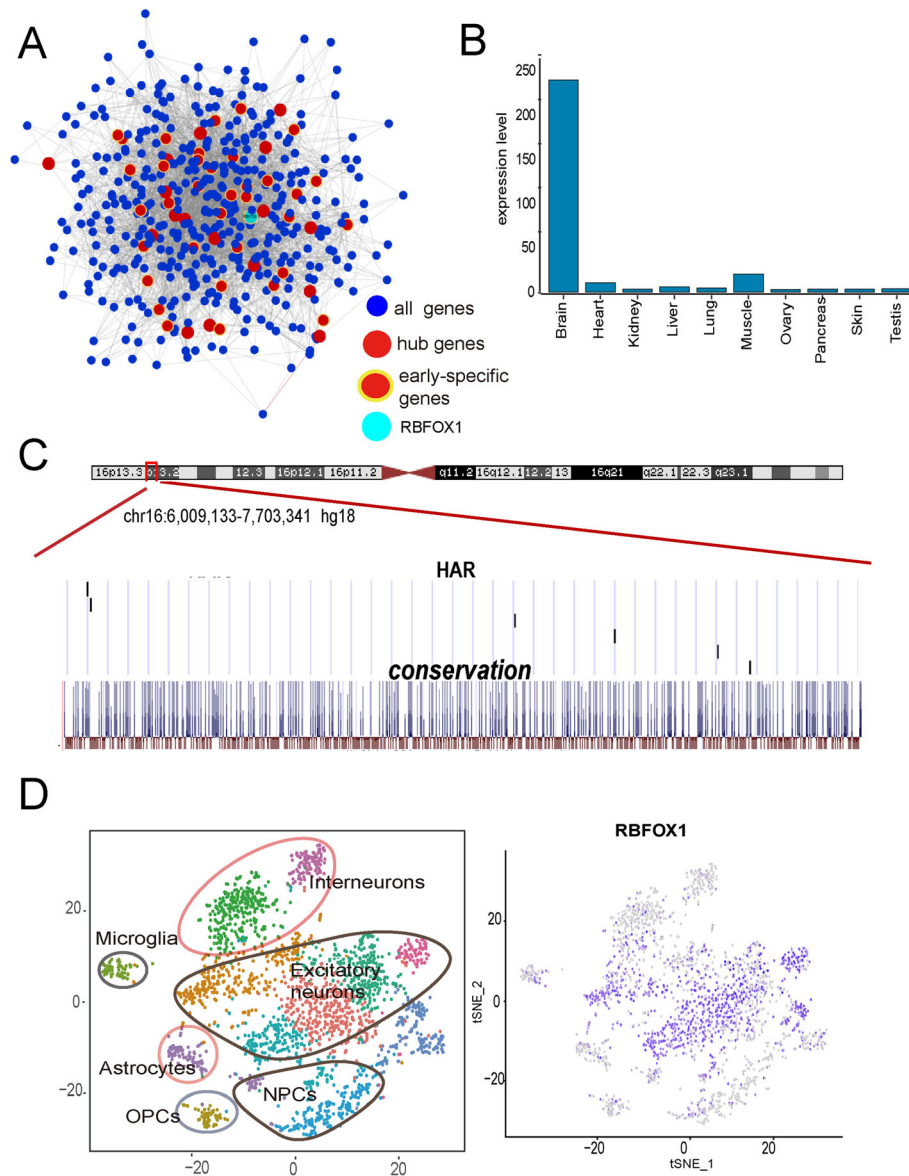


Fig. 6 Hub gene *RBFOX1* in module GCM1. **a** Network plot of hub genes identified within GCM1 module. Blue nodes indicate all genes. Red nodes indicate hub genes. Yellow halos indicate early stage-specific hub genes. Cyan node indicates *RBFOX1*. Edges reflect significant interactions between genes based on mutual information. **b** Expression level of *RBFOX1* in different tissues. **c** Location of HAEs at *RBFOX1* locus in human genome and conservation of *RBFOX1* among 17 mammals according to UCSC Genome Browser (www.genome.ucsc.edu). **d** Cell specificity of *RBFOX1*

gene expression profiles between humans and rhesus macaques, our results provide new insights into the gene expression trajectory of human brain development, which will deepen our understanding on evolution of the human brain.

Methods

Dataset resources

The normalized gene expression human and rhesus macaque datasets were obtained from the Allen Brain Atlas (<http://www.brain-map.org>) (Table 1; Table 2) [24, 25].

We used two datasets for humans, including RNA-seq and microarray data, which contained 14 brain regions and 27 developmental stages. The RNA-seq data were summarized to Gencode 10 gene-level reads per kilobase million mapped reads (RPKM), whereas the microarray data were based on the Affymetrix GeneChip Human Exon 1.0 ST Array platform. Several quality control measures were implemented to reduce errors due to spatial artifacts on the chips, technical differences between chips in probe saturation, or other unaccounted for batch effects. Detailed information can

be found in the Supplementary Materials of Kang et al. (2011) [24]. For rhesus macaques, we used the microarray dataset based on GeneChip Rhesus Macaque Genome Arrays from Affymetrix. From the 52,865 probe sets included in the macaque microarray data, 12,441 high-confidence probe sets were retained after quality control filtering. Detailed description of the macaque data can be found in Bakken et al. (2016) [25]. The macaque microarray dataset contained five brain regions (22 brain subregions) and nine developmental stages (Table 1; Table 2).

Clustering of genes in each tissue

The human microarray and RNA-seq datasets and rhesus macaque microarray dataset were used to cluster genes for each brain region according to their expression levels. To reduce the influence of technical noise, only genes with expression values of more than 0 in 80% of the available samples were considered expressed. Before clustering, we log₂ transformed and then z-transformed the expression levels (normalized the mean to 0 and variance to 1). Using agglomerative hierarchical clustering with the average and complete methods in the flash-Clust R package [62], the RNA-seq and microarray data from most human tissues were clustered into two groups, separated with a demarcation point of 25 PCW. The microarray data from the rhesus macaques were also clustered into two groups, with 17 PCW as the demarcation point.

Dynamic network biomarker (DNB) analysis

Based on DNB theoretical analysis [30, 32, 33], we can prove that when a system is near the critical state/transition phase, a dominant group of genes/molecules, i.e., DNBs, can drive transition of the dynamic process. These molecules must satisfy the following three criteria [30]:

- (1) Deviation (or fluctuation) for each molecule inside the dominant group (SD_d : standard deviation) drastically increases.
- (2) Correlation between molecules inside the dominant group (PCC_d : Pearson correlation coefficients in absolute values) rapidly increases.
- (3) Correlation between molecules inside and outside the dominant group (PCC_o : Pearson correlation coefficients in absolute values) rapidly decreases.

The dominant group is considered a DNB and plays an important role in phase transition. A quantification index (CI) considering all three criteria can then be used as the numerical signal of the critical state or transition phase and also for the identification of DNB members/molecule, with the following equation:

$$CI = \frac{SD_d \cdot PCC_d}{PCC_o} \quad (1)$$

where, PCC_d is the average Pearson's correlation coefficient (PCC) between the genes in the dominant group (or DNB) of the same time stage in absolute value; PCC_o is the average PCC between the dominant group (or DNB) and others of the same time stage in absolute value; and, SD_d is the average standard deviation (SD) of genes in the dominant group (or DNB). The three criteria together construct the composite index (CI). The CI is expected to peak or increase sharply during the measured stages when the system approaches the critical state, thus indicating imminent transition or transition phase of the biological process [30].

We applied this DNB method to detect the critical points and DNB members in humans and macaques based on the transcriptome data of the primary visual cortex. In each sampling stage, there were 1–4 samples with gene expression profiles. To increase the reliability of the DNB results, the slide window method was incorporated into the DNB model to process data [30].

Differentially expressed genes (DEGs) between early and later stages of brain development

To remove the potential effect of different high-throughput platforms on gene expression values, we only used microarray data for DEG analysis for both species. Pairwise differential expression was investigated using the edgeR R package [63]. To determine the DEGs between the two developmental times for humans and macaques, the demarcation times were set 25 PCW and 17 PCW, respectively. A nominal significance threshold of Benjamini-Hochberg FDR < 0.05 and fold change [FC] > 1.5 was used to identify DEGs.

For DEGs in humans, we applied g:Profiler (<https://biit.cs.ut.ee/gprofiler/>) [64] for functional annotation analysis (GO and KEGG). To assess cell-type specificity in the 14 brain regions of humans, we used genes expressed at least five-fold higher in one cell type than all other cell types (neuron, microglia, astrocyte, oligodendrocyte, endothelial) from brain-based RNA expression data as the cell marker [65].

Heterochrony analyses with dynamic time warping algorithm (DTW-S)

We combined data from microarray probes of humans and macaques to study heterochronic gene expression in five brain regions (i.e., hippocampus (HIP), striatum (STR), anterior cingulate cortex (ACC), amygdala (AMY) and primary visual cortex (VIC)). We used the “sva” R package [66] to remove batch effects between microarray datasets of humans and rhesus macaques.

To choose age-related genes, we first used a log₂ transformed age scale to ensure a more linear relationship between age and phenotype [67]. We tested the effect of age on the expression levels using polynomial regression models, as described previously [68]. We next tested each gene for expression divergence between humans and rhesus macaques using analysis of covariance [69] (F-test $P < 0.05$). Identification of age-related genes and species-specific genes were based on the adjusted r^2 criterion. The identification methods have been described previously [68]. Consequently, we selected 892 genes in AMY, 2431 genes in HIP, 1961 genes in STR, 1899 genes in ACC, and 2416 genes in VIC as the test gene set for DTW-S, satisfying the following criteria: (i) significant expression change with age and (ii) significant expression difference between humans and macaques.

The DTW-S algorithm was then used to analyze the data for heterochrony [68]. We defined genes showing significant heterochrony into two categories: (i) human acceleration genes, whose expression changes were significantly faster during human brain development than that during macaque brain development; and, (ii) human neoteny genes, whose expression changes were significantly delayed during human brain development compared with that during macaque brain development. Using those genes that showed significant age-related and species-specific differences, as defined above, we aligned the macaque and human expression trajectories and estimated the time-shift (heterochrony) between humans and macaques, with simulations conducted to estimate the significance of the shifts. We considered genes as 'significantly heterochronic' if they showed a shift at $P < 0.05$. A detailed description of the DTW-S algorithm can be found in Yuan et al. 2011 [68].

Construction of gene co-expression modules for early human brain development

We constructed multi-tissue co-expression networks that simultaneously captured intra- and inter-tissue gene-gene interactions using the human RNA-seq expression data [42, 70]. Before identifying co-expressed gene modules, we used linear regression to correct sex and brain region covariates in the expression data. To quantify the differences in the transcript network organization between the early and late stages, we employed the modular differential connectivity (MDC) metric [71]. In brief, MDC represents the connectivity ratios of all gene pairs in a module from the early stage to the same gene pairs from the later stage: $MDC > 0$ indicates a gain of connectivity or enhanced co-regulation between genes in the early stage, whereas $MDC < 0$

indicates a loss of connectivity or reduced co-regulation between genes in the early stage.

To identify key regulator (driver) genes of the GCM1 module, we applied key driver analysis to the module-based unweighted co-expression networks derived from ARACNE [47]. ARACNE first identified significant interactions between genes in the module based on their mutual information and then removed indirect interactions through data processing inequality (DPI). For each ARACNE-derived unweighted network, we further identified the key regulators by examining the number of N -hop neighborhood nodes for each gene.

Identification of cell type and subtype from single cell data

Single-cell RNA-seq data (accession number GSE104276) were reported in previous study [39]. Transcript counts for each cell were normalized to transcripts per million (TPM), with the TPM values then normalized by $\log((TPM/10)+1)$ for subsequent analysis [39]. The Seurat package [72] v1.2.1 implemented in R was applied to identify major cell types among the 2394 single cells from the prefrontal cortex. Only cells that expressed more than 1000 genes were considered, and only genes with normalized expression levels greater than 1 and expressed in at least three single cells were included, which left 20,262 genes across 2344 samples for clustering analysis. After initial clustering, *PAX6*, *NEUROD2*, *GAD1*, *PDGFRA*, *AQP4*, and *PTPRC* were used as markers to identify the major cell types in the brain: i.e., neural progenitor cells, excitatory neurons, interneurons, oligodendrocyte progenitor cells, astrocytes, and microglia, respectively.

Coding sequence evolutionary analysis of *RBFOX1*

To analyze the evolution of the coding regions of *RBFOX1*, we obtained the human, chimpanzee, rhesus macaque, marmoset, mouse lemur, mouse, rat, cow, dog, and opossum coding sequences for this gene from Ensembl [48]. The coding sequences were aligned using Prank [73]. Gblocks v0.91b was used to remove poorly aligned regions in the resulting nucleotide sequence alignments [74]. We then used the modified branch-site model A from the PAML package v4.9 to test positive selection of *RBFOX1* in the human and primate lineages, respectively [75]. The null hypothesis of the branch test was that all lineages shared the same dN/dS ratio. The alternative hypothesis was that human or primate lineages had a different dN/dS ratio from other lineages, with w_0 , w_1 , and w_2 representing codons under negative, null, and positive selection, respectively. The Chi-square test was used to calculate the P value. $P_{\text{adjust}} < 0.05$ was considered as significant.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12862-020-01633-4>.

Additional file 1: Supplementary Tables S1–S12.

Additional file 2: Supplementary Figure S1–S5.

Abbreviations

PCW: Postconception week; DNB : Dynamic network biomarker; NHPS : Non-human primates; RNA-seq : RNA-sequencing; HIP : Hippocampus; STR: Striatum; ACC : Anterior cingulate cortex; AMY : Amygdala; V1C: Primary visual cortex; DEGs : Differentially expressed genes; FC: Fold change; GW: Gestational week; WGCNA : Gene co-expression network analysis; MDC: Modular differential connectivity; GCM: Modules that show gain of connectivity; LCM : Modules that show loss of connectivity; NCM : Modules that show no change of connectivity; *RBFOX1*: RNA Binding Fox-1 Homolog 1; HACNSs : Human-accelerated conserved non-coding sequences; HARs: Human DNA sequence accelerated regions; TPM : Transcripts per million

Acknowledgements

This work was supported by the Animal Branch of the Germplasm Bank of Wild Species, Chinese Academy of Sciences (the Large Research Infrastructure Funding).

Authors' contributions

D.D.W. led the project. D.D.W., Y.L., and L.N.C. designed and conceived the study. L.M.L. performed primary analyses. H. T conducted DNB analysis. L.M.L.Y.S. and M.S.W. contributed to the positive selection analysis of *RBFOX1*. L.M.L.H.B.X. and S.W. contributed to single cell analysis. D.D.W., L.M.L., D.M.I., and A.A. prepared the paper. All authors read and approved the manuscript.

Funding

This work was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB13000000, XDB38040400), National Key R&D Program of China (2017YFA0505500), National Natural Science Foundation of China (31671325, 31822048, 31771476, 31930022), and the Bureau of Science and Technology of Yunnan Province (2019FY1010). The funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript. Y.L. was supported by the Young Academic and Technical Leader Raising Foundation of Yunnan Province.

Availability of data and materials

The RNA-seq and microarray dataset of human analysed during the current study are available in the <http://www.brainspan.org/static/download.html> repository [24]. The microarray dataset of macaque analysed during the current study are available in the <http://blueprintpatlas.org/static/download-repository> [25].

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, Yunnan, China.

²Kunming College of Life Science, University of the Chinese Academy of Sciences, Kunming 650223, Yunnan, China. ³State Key Laboratory of Cell Biology, Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Shanghai 200031, China.

⁴Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario M5S 1A8, Canada. ⁵Banting and Best Diabetes Centre, University of Toronto, Toronto, Ontario M5G 2C4, Canada. ⁶Key

Laboratory of Systems Biology, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Hangzhou 310024, China. ⁷Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, Yunnan, China. ⁸State Key Laboratory for Conservation and Utilization of Bio-Resource, Yunnan University, Kunming 650091, Yunnan, China.

Received: 16 April 2019 Accepted: 26 May 2020

Published online: 23 June 2020

References

- Sousa AMM, Meyer KA, Santpere G, Gulden FO, Sestan N. Evolution of the human nervous system function, structure, and development. *Cell*. 2017; 170(2):226–47.
- Workman AD, Charvet CJ, Clancy B, Darlington RB, Finlay BL. Modeling transformations of neurodevelopmental sequences across mammalian species. *J Neurosci*. 2013;33(17):7368.
- Bae B-I, Jayaraman D, Walsh CA. Genetic changes shaping the human brain. *Dev Cell*. 2015;32(4):423–34.
- Lieberman P. The evolution of language and thought. *J Anthropol Sci*. 2016; 94:127–46.
- Penn DC, Holyoak KJ, Povinelli DJ. Darwin's mistake: explaining the discontinuity between human and nonhuman minds. *Behav Brain Sci*. 2008; 31(2):109–30 discussion 130–178.
- Zhang M-L. M-LLAOARWMD-DWYS: Conserved sequences identify the closest living relatives of primates. *Zool Res*. 2019;40(6):532–40.
- Enard W. The molecular basis of human Brain evolution. *Curr Biol*. 2016; 26(20):R1109–17.
- Franchini LS, Pollard K. Genomic approaches to studying human-specific developmental traits. *Development*. 2015;142:3100–12.
- Silver DL. Genomic divergence and brain evolution: how regulatory DNA influences development of the cerebral cortex. *BioEssays*. 2016;38(2):162–71.
- Chugani HT, Phelps ME, Mazziotta JC. Positron emission tomography study of human brain functional development. *Ann Neurol*. 1987;22(4):487–97.
- Peter RH. Synaptic density in human frontal cortex — developmental changes and effects of aging. *Brain Res*. 1979;163(2):195–205.
- Sterner K, Weckle AT, Chugani H, Tarca A, Sherwood C, Hof P, Kuzawa C, Boddy A, Abbas A, Raaum R, et al. Dynamic Gene Expression in the Human Cerebral Cortex Distinguishes Children from Adults. *PLoS one*. 2012;7:e37714.
- Jacobs B, Chugani HT, Allada V, Chen S, Phelps ME, Pollack DB, Raleigh MJ. Developmental changes in Brain metabolism in sedated rhesus macaques and Vervet monkeys revealed by positron emission tomography. *Cereb Cortex*. 1995;5(3):222–33.
- Ye L-Q, Zhao H, Zhou H-J, Ren X-D, Liu L-L, Otecko NO, Wang Z-B, Yang M-M, Zeng L, Hu X-T, et al. The RNA editome of Macaca mulatta and functional characterization of RNA editing in mitochondria. *Sci Bull*. 2017; 62(12):820–30.
- Cáceres M, Lachuer J, Zapala MA, Redmond JC, Kudo L, Geschwind DH, Lockhart DJ, Preuss TM, Barlow C. Elevated gene expression levels distinguish human from non-human primate brains. *Proc Natl Acad Sci*. 2003;100(22):13030.
- Gu J, Gu X. Induced gene expression in human brain after the split from chimpanzee. *Trends Genet*. 2003;19(2):63–5.
- Kim J, Kerr JQ, Min G-S. Molecular heterochrony in the early development of *Drosophila* and *em*. *Proc Natl Acad Sci*. 2000;97(1):212.
- Langer J. The Heterochronic evolution of primate cognitive development. *Biol Theory*. 2006;1(1):41–3.
- Moss EG. Heterochronic genes and the nature of developmental time. *Curr Biol*. 2007;17(11):R425–34.
- Somel M, Franz H, Yan Z, Lorenc A, Guo S, Giger T, Kelso J, Nickel B, Dannemann M, Bahn S, et al. Transcriptional neoteny in the human brain. *Proc Natl Acad Sci U S A*. 2009;106(14):5743–8.
- Stiles J, Jernigan TL. The basics of brain development. *Neuropsychol Rev*. 2010;20(4):327–48.
- Zhu Y, Sousa AMM, Gao T, Skarica M, Li M, Santpere G, Esteller-Cucala P, Juan D, Ferrández-Peral L, Gulden FO, et al. Spatiotemporal transcriptomic divergence across human and macaque brain development. *Science*. 2018; 362(6420):eaat8077.
- Khaitovich P, Muetzel B, She X, Lachmann M, Hellmann I, Dietzsch J, Steigele S, Do H-H, Weiss G, Enard W, et al. Regional patterns of gene

- expression in human and chimpanzee brains. *Genome Res.* 2004;14(8):1462–73.
24. Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M, Sousa AM, Pletikos M, Meyer KA, Sedmak G, et al. Spatio-temporal transcriptome of the human brain. *Nature.* 2011;478(7370):483–9.
 25. Bakken TE, Miller JA, Ding S-L, Sunkin SM, Smith KA, Ng L, Szafer A, Dalley RA, Royall JJ, Lemon T, et al. A comprehensive transcriptional map of primate brain development. *Nature.* 2016;535:367.
 26. Sunkin SM, Ng L, Lau C, Dolbeare T, Gilbert TL, Thompson CL, Hawrylycz M, Dang C. Allen Brain atlas: an integrated spatio-temporal portal for exploring the central nervous system. *Nucleic Acids Res.* 2013;41(Database issue):D996–D1008.
 27. Li M-L, Wu S-H, Zhang J-J, Tian H-Y, Shao Y, Wang Z-B, Irwin DM, Li J-L, Hu X-T, Wu D-D. 547 transcriptomes from 44 brain areas reveal features of the aging brain in non-human primates. *Genome Biol.* 2019;20(1):258.
 28. Raznahan A, Greenstein D, Lee NR, Clasen LS, Giedd JN. Prenatal growth in humans and postnatal brain maturation into late adolescence. *Proc Natl Acad Sci.* 2012;109(28):11366.
 29. Somel M, Guo S, Fu N, Yan Z, Hu HY, Xu Y, Yuan Y, Ning Z, Hu Y, Menzel C, et al. MicroRNA, mRNA, and protein expression link development and aging in human and macaque brain. *Genome Res.* 2010;20(9):1207–18.
 30. Chen L, Liu R, Liu ZP, Li M, Aihara K. Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers. *Sci Rep.* 2012;2:342.
 31. Richard A, Boullu L, Herbach U, Bonnafoux A, Morin V, Vallin E, Guillemin A, Papiili Gao N, Gunawan R, Cosette J, et al. Single-cell-based analysis highlights a surge in cell-to-cell molecular variability preceding irreversible commitment in a differentiation process. *PLoS Biol.* 2016;14(12):e1002585.
 32. Liu X, Chang X, Liu R, Yu X, Chen L, Aihara K. Quantifying critical states of complex diseases using single-sample dynamic network biomarkers. *PLoS Comput Biol.* 2017;13(7):e1005633.
 33. Liu R, Wang X, Aihara K, Chen L. Early diagnosis of complex diseases by molecular biomarkers, network biomarkers, and dynamical network biomarkers. *Med Res Rev.* 2014;34(3):455–78.
 34. Cahoy JD, Emery B, Kaushal A, Foo LC, Zamanian JL, Christopherson KS, Xing Y, Lubischer JL, Krieg PA, Krupenko SA, et al. A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *J Neurosci.* 2008;28(1):264–78.
 35. Silbereis JC, Pochareddy S, Zhu Y, Li M, Sestan N. The cellular and molecular landscapes of the developing human central nervous system. *Neuron.* 2016; 89(2):248–68.
 36. Shen Q, Wang Y, Dimos JT, Fasano CA, Phoenix TN, Lemischka IR, Ivanova NB, Stifani S, Morrissy EE, Temple S. The timing of cortical neurogenesis is encoded within lineages of individual progenitor cells. *Nat Neurosci.* 2006; 9(6):743–51.
 37. Qian X, Shen Q, Goderie SK, He W, Capela A, Davis AA, Temple S. Timing of CNS cell generation: a programmed sequence of neuron and glial cell production from isolated murine cortical stem cells. *Neuron.* 2000;28(1):69–80.
 38. Miller FD, Gauthier AS. Timing is everything: making neurons versus glia in the developing cortex. *Neuron.* 2007;54(3):357–69.
 39. Zhong S, Zhang S, Fan X, Wu Q, Yan L, Dong J, Zhang H, Li L, Sun L, Pan N, et al. A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex. *Nature.* 2018;555(7697):524–8.
 40. Li X, Newbern JM, Wu Y, Morgan-Smith M, Zhong J, Charron J, Snider WD. MEK is a key regulator of Gliogenesis in the developing Brain. *Neuron.* 2012; 75(6):1035–50.
 41. Bagot RC, Cates HM, Purushothaman I, Lorsch ZS, Walker DM, Wang J, Huang X, Schluter OM, Maze I, Pena CJ, et al. Circuit-wide transcriptional profiling reveals Brain region-specific gene networks regulating depression susceptibility. *Neuron.* 2016;90(5):969–83.
 42. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics.* 2008;9:559.
 43. Parikshak NN, Luo R, Zhang A, Won H, Lowe JK, Chandran V, Horvath S, Geschwind DH. Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell.* 2013;155(5):1008–21.
 44. Meyer KA, Marques-Bonet T, Sestan N. Differential gene expression in the human Brain is associated with conserved, but not accelerated, noncoding sequences. *Mol Biol Evol.* 2017;34(5):1217–29.
 45. Lindblad-Toh K, Garber M, Zuk O, Lin MF, Parker BJ, Washietl S, Kheradpour P, Ernst J, Jordan G, Maudeli E, et al. A high-resolution map of human evolutionary constraint using 29 mammals. *Nature.* 2011;478(7370):476–82.
 46. Willsey AJ, Sanders SJ, Li M, Dong S, Tebbenkamp AT, Muhle RA, Reilly SK, Lin L, Fertuzinhos S, Miller JA, et al. Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell.* 2013;155(5):997–1007.
 47. Margolin AA, Nemenman I, Basso K, Wiggins C, Stolovitzky G, Dalla Favera R, Califano A. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC bioinformatics.* 2006;7(Suppl 1):S7.
 48. Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, Billis K, Cummins C, Gall A, Girón CG, et al. Ensembl 2018. *Nucleic Acids Res.* 2018; 46(D1):D754–61.
 49. Alkallas R, Fish L, Goodarzi H, Najafabadi HS. Inference of RNA decay rate from transcriptional profiling highlights the regulatory programs of Alzheimer's disease. *Nat Commun.* 2017;8(1):909.
 50. Lal D, Reinthaler EM, Altmüller J, Toliat MR, Thiele H, Nurnberg P, Lerche H, Hahn A, Moller RS, Muhle H, et al. RFXO1 and RFXO3 mutations in rolandic epilepsy. *PLoS One.* 2013;8(9):e73323.
 51. Lee JA, Damianov A, Lin CH, Fontes M, Parikshak NN, Anderson ES, Geschwind DH, Black DL, Martin KC. Cytoplasmic Rbfox1 regulates the expression of synaptic and autism-related genes. *Neuron.* 2016;89(1):113–28.
 52. Gehman LT, Stoilov P, Maguire J, Damianov A, Lin CH, Shiue L, Ares M Jr, Mody I, Black DL. The splicing regulator Rbfox1 (A2BP1) controls neuronal excitation in the mammalian brain. *Nat Genet.* 2011;43(7):706–11.
 53. Kamm GB, Pisciotto F, Kliger R, Franchini LF. The developmental brain gene NPAS3 contains the largest number of accelerated regulatory sequences in the human genome. *Mol Biol Evol.* 2013;30(5):1088–102.
 54. He Z, Han D, Efimova O, Gujjarro P, Yu Q, Oleksiak A, Jiang S, Anokhin K, Velichkovsky B, Grünewald S, et al. Comprehensive transcriptome analysis of neocortical layers in humans, chimpanzees and macaques. *Nat Neurosci.* 2017;20:886.
 55. Kronenberg ZN, Fiddes IT, Gordon D, Murali S, Cantsilieris S, Meyerson OS, Underwood JG, Nelson BJ, Chaisson MJP, Dougherty ML, et al. High-resolution comparative analysis of great ape genomes. *Science.* 2018;360: eaar6343.
 56. Sousa AMM, Zhu Y, Raghanti MA, Kitchen RR, Onorati M, Tebbenkamp ATN, Stutz B, Meyer KA, Li M, Kawasawa YI, et al. Molecular and cellular reorganization of neural circuits in the human lineage. *Science.* 2017; 358(6366):1027–32.
 57. Vermunt MW, Tan SC, Castelijn B, Geeven G, Reinink P, de Bruijn E, Kondova I, Persengiev S, Netherlands Brain B, Bontrop R, et al. Epigenomic annotation of gene regulatory alterations during evolution of the primate brain. *Nat Neurosci.* 2016;19:494.
 58. Xu C, Li Q, Efimova O, He L, Tatsumoto S, Stepanova V, Oishi T, Udono T, Yamaguchi K, Shigenobu S, et al. Human-specific features of spatial gene expression and regulation in eight brain regions. *Genome Res.* 2018;28(8):1097–110.
 59. Bakken TE, Miller JA, Luo R, Bernard A, Bennett JL, Lee C-K, Bertagnolli D, Parikshak NN, Smith KA, Sunkin SM, et al. Spatiotemporal dynamics of the postnatal developing primate brain transcriptome. *Hum Mol Genet.* 2015; 24(15):4327–39.
 60. Bianchi S, Stimpson CD, Bauernfeind AL, Schapiro SJ, Baze WB, McArthur MJ, Bronson E, Hopkins WD, Semendeferi K, Jacobs B, et al. Dendritic morphology of pyramidal neurons in the chimpanzee neocortex: regional specializations and comparison to humans. *Cereb Cortex.* 2013;23(10):2429–36.
 61. Elston G, Benavides-Piccione R, Elston A, Manger P, Defelipe J. Pyramidal cells in prefrontal cortex of Primates: marked differences in neuronal structure among species. *Front Neuroanat.* 2011;5:2.
 62. Langfelder P, Horvath S. Fast R functions for robust correlations and hierarchical clustering. *J Stat Softw.* 2012;46(11):i11.
 63. Guttman M, Garber M, Levin JZ, Donaghey J, Robinson J, Adiconis X, Fan L, Koziol MJ, Gnirke A, Nusbaum C, et al. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nat Biotechnol.* 2010;28(5):503–10.
 64. Reimand J, Arak T, Adler P, Kolberg L, Reisberg S, Peterson H, Vilo J. g: profiler—a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Res.* 2016;44(W1):W83–9.
 65. Zhang Y, Chen K, Sloan SA, Bennett ML, Scholze AR, O'Keefe S, Phatnani HP, Guarnieri P, Caneda C, Ruderis N, et al. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J Neurosci.* 2014;34(36):11929–47.
 66. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics.* 2012;28(6):882–3.

67. Clancy B, Darlington RB, Finlay BL. Translating developmental time across mammalian species. *Neuroscience*. 2001;105(1):7–17.
68. Yuan Y, Chen Y-PP, Ni S, Xu AG, Tang L, Vingron M, Somel M, Khaitovich P. Development and application of a modified dynamic time warping algorithm (DTW-S) to analyses of primate brain expression time series. *BMC bioinformatics*. 2011;12:347.
69. Faraway JJ. *Practical regression and ANOVA using R*; 2002.
70. Zhang B, Gaiteri C, Bodea LG, Wang Z, McElwee J, Podtelezchnikov AA, Zhang C, Xie T, Tran L, Dobrin R, et al. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell*. 2013;153(3):707–20.
71. McKenzie AT, Katsyv I, Song WM, Wang M, Zhang B. DGCA: a comprehensive R package for differential gene correlation analysis. *BMC Syst Biol*. 2016;10(1):106.
72. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018;36(5):411–20.
73. Loytynoja A. Phylogeny-aware alignment with PRANK. *Methods Mol Biol*. 2014;1079:155–70.
74. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol*. 2000;17(4):540–52.
75. Yang Z, dos Reis M. Statistical properties of the branch-site test of positive selection. *Mol Biol Evol*. 2011;28(3):1217–28.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

