Single-cell transcriptomic analysis identifies neocortical developmental differences between human and mouse

Ziheng Zhou[1,2], Shuguang Wang[1,2,3], Dengwei Zhang[1,2], Xiaosen Jiang[1,2,4], Jie Li[1,2,4], Ying Gu[1,2,6*], Hai-Xi Sun[1,2,6*]

[1] BGI-Shenzhen, Shenzhen 518083, China

[2] China National GeneBank, BGI-Shenzhen, Shenzhen 518120, China

[3] College of Informatics, Huazhong Agricultural University, Wuhan, 430070 Hubei P.R. China

[4] BGI Education Center, University of Chinese Academy of Sciences, Shenzhen 518083, China

[5] Institute for Stem cell and Regeneration, Chinese Academy of Sciences, Beijing, 100101, P. R. China

[*] Corresponding authors: sunhaixi@genomics.cn (H.-X.S.), guying@genomics.cn (Y.G.)

Abstract

Background: The specification and differentiation of neocortical projection neurons is a complex process under precise molecular regulation; however, little is known about the similarities and differences in cerebral cortex development between human and mouse at single-cell resolution.

Results: Here, using single-cell RNA-seq (scRNA-seq) data we explore the divergence and conservation of human and mouse cerebral cortex development using 18,446 and 7,610 neocortical cells. Systematic cross-species comparison reveals that the overall transcriptome profile in human cerebral cortex is similar to that in mouse such as cell types and their markers genes. By single-cell trajectories analysis we find human and mouse excitatory neurons have different developmental trajectories of neocortical projection neurons, ligand-receptor interactions and gene expression patterns. Further analysis reveals a refinement of neuron differentiation that occurred in human but not in mouse, suggesting that excitatory neurons in human undergo refined transcriptional states in later development stage. By contrast, for glial cells and inhibitory neurons we detected conserved developmental trajectories in human and mouse.

Conclusions: Taken together, our study integrates scRNA-seq data of cerebral cortex development in human and mouse, and uncovers distinct developing models in neocortical projection neurons. The earlier activation of cognition-related genes in human may explain the differences in behavior, learning or memory abilities between the two species.

Keywords:

Cerebral cortex development, Cross-species clustering, Neocortical projection neurons, Cognition

Background

The mammalian cerebral cortex develops via a complex process of cell proliferation, differentiation, and migration events. The molecular features of the human brain at gestational weeks 7-23 are similar to those of the mouse at postnatal days 14.5 to birth (P0) and reveal gene expression differences between the two species [1, 2]. Interneurons and projection neurons, which are the two major classes of neurons, populate the neocortex. Interneurons show largely GABAergic, connecting locally within the neocortex, and are generated by progenitors in the ventral telencephalic (inhibitory cortical interneuron producing) radial glia, before migrating from sub pallial ganglionic eminences into the cortex [3-6]. By contrast, projection neurons are excitatory, which are generated by progenitors in the dorsal telencephalic (excitatory cortical neuron producing) radial glia [7-9].

Despite immense efforts over the past decades to study the cell types, function, differentiation for the mammalian cerebral cortex, especially human and mouse, large differences and conservation are observed in human versus animal models [10-12]. Elucidating the conserved, divergent and cellular architecture of developing the cerebral cortex, especially projection neurons developments, may highlight to understanding susceptibility to neurological diseases and the occurrence of cognition and memory abilities in the early stages.

Recently, with the advance of next-generation sequencing (NGS) techniques, single-cell RNA sequencing (scRNA-seq) technologies has been paved the way for study cell phenotype and cell behavior at single-cell resolution [13, 14]. More significantly, scRNA-seq can overcome the heterogeneity of cross-species analysis, bypassing the need for pure cell sorting strategies [15, 16]. Moreover, although many studies have been reported the single species cerebral cortex using scRNA-seq, there is no systematic research to

integrate the human and mouse scRNA-seq data, and compare the conservation and divergence of developing neocortex.

Here we comprehensively characterize the conservation and divergence of the cerebral cortex program between human and mouse [9, 17, 18]. We present a cross-species cell atlas and in which no obvious differences in cell type composition was observed. However, we identified different developing trajectories of projection neurons that in human the subventricular zone (SVZ) migration to deeper neocortical layers and more superficial layers are two independent processes. Additionally, the deeper and upper layers are more precisely differentiated in human than those in mouse, and lack of neuron migration signal in human. By comparing excitatory neurons gene expression across species, we also observed significant expression differences of function genes for cell migration, cognition, learning and memory in humans compared with mice. We next evaluated whether further refinement of neuronal subtypes occurs as neurons mature, indicating that human neuronal subtypes resolve from coarse laminar distinctions to refined transcriptional states in the later development, whereas mouse may not show precise subtype differentiation. In summary, we show the importance of cross-evolutionary comparison to better characterize the differences of human and mouse projection neurons development, and the relevance of these to cognition pathway. In conclusion, our study offers essential clues for researching neurodevelopment, especially for the projection neurons development, and the differences of occurrence of cognition ability in human and mouse.

RESULTS

Integration of single-cell transcriptomics of the human and mouse cerebral cortex

In order to compare cellular and molecular processes of the late corticogenesis between human and mouse, we collected published scRNA-seq data sets of human and mouse cerebral cortex at different development stages [9, 17, 18]. We reanalyzed the transcriptome profiles of 18,446 mouse neocortical cells at two key time points for corticogenesis development including 10,836 cells at embryonic day 14.5 (E14.5) from six biological replicates (Fig. 1A), and 7,610 cells at birth (P0) from three biological replicates (Fig. 1A). Because 7- to 23-post-conception-week (pcw) human brain cells are analogous to E14.5 to P0 mouse brain cells (Fig. 1A) [1, 2], we incorporated the transcriptome profiles of 6,655 human brain cells including 100 cells between 5.85 and 7 pcw, and 800 cells between 23-37 pcw (Fig. 1A). To ensure that cells from different species and data sets can be properly integrated, we performed unbiased clustering using Uniform Manifold Approximation and Projection (UMAP) [19]. The major cell lineages in cerebral cortex such as excitatory neurons, inhibitory neurons, neuronal progenitor cells and glial cells were able to be identified using their lineage markers [20, 21] (Fig. 1B, C). Moreover, there was no obvious bias because the majority of the cell types could be identified in all three data sets (Fig. S1C).

Because the number of mouse cells (18,446) was about 3 times higher than that of human cells (6,655) which may cause bias, we randomly extracted 6,655 mouse cells in further analysis. Finally, 13,278 cells (6,655 mouse cells and 6,623 human cells) were used (Fig. 1A) in further analysis which covers all major cell types (Fig. 1D, E).

Identification and characterization of cortical cell types

We first compared the cell composition and found that mouse had more interneuron than human (Fig. 2A). We next compared the cell number in each cluster and found that cluster 06 Inh_striatal and cluster 09 Layer [19P] had more cells in mouse (20%-30%) than in human (70%-80%). In contrast, cluster 16 Microglia and 17 Mesenchyme had more human cells (80%-90%) (Fig. 2B).

We next examined the correlation among clusters by cross-species clustering (Fig. 2C). As there were still considerable species differences between humans and mice, we used the integrated matrix. Remarkably, excitatory neuron exhibited a profile highly similar to interneuron, and Glia showed itself a high degree of correlation (Fig. 2C). Interestingly, Cluster 16 Microglia were distinct from excitatory neuron, interneuron and other Glia cells (Fig. 2C). We also calculated the correlation using human and mouse cells respectively and found that human and mouse had similar results with cross-species (Fig. S2A, B).

We checked the top 10 highly expressed genes of each cluster and found most of them were consistent among human and mouse (Fig. 2D, S2C, D). For example, two SVZ clusters (Clusters 00 SVZ and 01 SVZ) were featured by predominant expression of *EOMES* (*TBR2*) and *ZBTB20* [22-24]. Cells in Cluster 01 SVZ also expressed markers of both excitatory (*NEUROD6*) and inhibitory (*CALB2*) neurons [25-27] which exhibited the expression patterns of rostral migratory steam [18]. Cells in deep-layer populations (Clusters 02 Exc deep and 03 Exc deep) exhibited high expression levels of *BCL11B* [28]. Cells in Clusters 04 Exc upper and 05 Exc upper expressed upper layer marker *MEF2C* [8, 29]. We identified *NEUROD6* and *NEUROD1* kept consistently high expression in excitatory inhibitory, and *EOMES* expressed in subventricular zone (SVZ), and *STMN2* expressed strongly in deep and upper layer and increased the levels of expression with neuron maturation, which conversed the transformation known genes expression trajectories in neurogenesis among human and mouse (from *EOMES* to *STMN2*) [9, 30]

(Fig. 2D). Three interneuron clusters (Clusters 06 Inh_striatal, 07 Inh and 08 Inh) were characterized by expression of *GAD2* and *NRXN3*. Specifically, in Cluster 08 Inh we found high expression of *LHX6*, *SST* and *MAF* indicating 08 Inh is somatostatin (SST) expressing interneurons but not parvalbumin (PV). *LHX6* is a transcription factor associated with PV and SST interneuron [31, 32], whereas *SST* and *MAF* regulate the potential of interneurons to interneurons to acquire SST+ interneuron identity [33]. Interneuron migrate tangentially from the ganglionic eminences, and ganglionic eminences (Cluster 13 GE) expressed high levels of inhibitory marker *DLX1* and proliferation markers *TOP2A* and *MKI67* [34]. RG marker *HES5* was in two RG clusters (Clusters 10 RG and 11 RG), while Cluster 11 RG expressed high levels of *MKI67* and *TOP2A* genes. We also found *GAP43*, a key regulator in early brain development, was constitutively expressed in all these cell types in both human and mouse [35].

We identified Clusters 06 Inh_striatal and 09 Layer[19P] as inhibitory neurons and they had similar gene expression patterns and cell numbers in human and mouse. Striatal inhibitory neurons (Cluster 06 Inh_striatal) expressed inhibitory markers *GAD2* and lower-layer marker *BCL11B* [36]. Layer 19-P (Cluster 09 layer[19P]) expressed canonical Cajal-Retzius cell markers *RELN* and Purkinje cell marker *PCP4*. In addition, they also expressed non-coding RNAs *SNHG11* and *MEG3* (Fig. 2D, S2C). Microglia (Cluster 16 Microglia) had higher cell numbers in human than in mouse and exhibited low expression correlation with other clusters. They expressed microglial markers such as *LAPTM5*, *C1QA*, *C1QB*, and *C1QC* [37], as well as several microglia-specific pathways such as TREM2-TYROBP pathway [37-40] (Fig. 2D). Interestingly, Cluster 20 Unknown exhibited an intermediate state between excitatory neuron and interneuron: it had high expression correlation with both excitatory neuron and interneuron, and expressed several marker genes such as *STMN2* (excitatory neuron) and *NNAT* (interneuron) (Fig. 2D, S2C, D).

Different developmental trajectories of excitatory neuron between human and

mouse

To compare the developmental trajectories in excitatory neuron between human and mouse, we split the data into 6,623 human cells and 6,655 mouse cells, and performed unbiased clustering using UMAP again. We identified 23 clusters and also observed lineage separation of the major cell types (Fig. S3A, B, C, D). We used *NEUROD6* as marker gene and identified 6 excitatory neuron clusters in human (Cluster 00 SVZ, 01 Exc deep, 02 Exc deep, 03 Exc upper, 04 Exc upper and 05 Exc upper) and 7 excitatory neuron clusters in mouse (Cluster 00 SVZ, 01 SVZ, 02 Exc deep, 03 Exc deep, 04 Exc deep, 05 Exc upper and 06 Exc upper) (Fig. S3A, B). We then reconstructed the corticogenesis developmental trajectories using Monocle analysis. Both human cells and mouse cells were distributed along pseudo-temporally ordered paths from SVZ to neocortical layers, including deeper neocortical layers (layer VI, then layer V) and superficial layers (layer IV, then layer II/III) [28] (Fig. 3A, B). Surprisingly, we found that the developmental trajectory in mouse was a continuous process without any branch, in which SVZ migrated to deeper neocortical layers and then to superficial layers. By contrast, in human this process had two branches in which SVZ migrated to both deeper neocortical layers and superficial layers (Fig. 3A, B, S3E, F). Consistent with this observation, we found that SVZ marker *EOMES* was specifically expressed in the beginning of trajectories, deep-layer marker *BCL11B* was specifically expressed in the upper branch in human, whereas migratory and upper-layer markers *POU3F2* and *SATB2* were specifically expressed in the lower branch, which was different from those in mouse trajectory (Fig. 3A, B, S3I). More important, when we further split the data based on development stages, we still found that the lower-layer neurons were present in 7 to 23 pcw in human and E14.5 and P0 in mouse, with high expression of *BCL11B*. Meanwhile, a large number of Layer II-IV (upper layer) cells at P0 in mouse and 20-23 pcw in human was identified, with high expression of *SATB2* (Fig. S3G, H).

We identified 109 differentially expressed genes in human that may determine cell fate commitment. They were fall into 4 clusters which exhibited successive waves of gene expression in different branches (Fig. 3E). In addition, Gene Ontology (GO) enrichment analysis revealed that the functions of genes predominantly expressed at the SVZ branch were enriched in cerebral cortex development, olfactory bulb development and olfactory lobe development as previously reported [41]. Among these genes, transcription factor *PAX6* is crucial for SVZ neurogenesis and also acts via retinoid signaling to regulate eye development [42] suggesting that eye development may have relative to SVZ neurogenesis (Fig. 3E). Lower-layer from human cortex showed increased expression of genes for modulation of synaptic signaling and chemical synaptic transmission, which have a mechanistic link with learning and memory by plasticity regulation [43]; whereas upper layer CPN from human cortex exhibited increased expression of genes involved in regulation of sodium ion transport and calcium-mediated signaling (Fig. 3E). Genes related to forebrain neuron fate commitment and forebrain neuron differentiation were expressed in both lower-layer and upper layer. More interestingly, the functions of genes that were highly expressed in SVZ, lower layer and upper layer were all enriched in cognition and learning (Fig. 3E).

We further analyzed the migratory marker genes expression in neocortical. We found these genes had continuously increased expression in SVZ, and decreased expression in deep-layer and upper-layer in human. Specifically, *NRP1* and *PALMD*, which were involved in the regulation of cell morphogenesis and cerebral cortex cell migration, exhibited increased expression in the SVZ lineage and then declined in deep layer and upper layer lineage, suggesting that neuronal migration in projection neurons in the cerebral cortex may occur from the SVZ's maturity stage to the early development of deep and shallow layers, rather than the late deep layer [44] (Fig. S3J, K). By contrast, in mouse

cerebral cortex we found migratory genes showed increased expression in SVZ and deep layer, and decreased expression in upper layer suggesting that in mouse the deep layer continuously process cell migration from SVZ to upper layer. Representative genes include *Meis2*, *Nrp1* and *Pou3f2*, whose functions are relative to cell migration and regulation of cell morphogenesis involved in differentiation (Fig. S3K).

Divergence in neocortical between human and mouse

In order to better understand the distinct layers in neocortical in human and mouse, we performed a comparison of cortical expression models through single-cell analysis, including deeper neocortical cortex (layer VI, then layer V) and late-born neurons migrate through deep layers to gradually fill more shallow layers (layer IV, then layer II/III). Pairwise comparison of human and mouse cortical layer data sets identified a large number of differentially expressed genes (DEGs) (Fig. 4A). We observed a considerably bigger number of differentially expressed genes in human than in mouse, meanwhile we also found some co-expression marker genes, for example, fez family zinc finger 2 *Fezf2* (high expressed in deeper layers), special AT-rich sequence binding protein 2 (*SATB2*) and *POUEF2*, a marker of layer II/III with function of migration and differentiation [28]. More interestingly, we detected the differentially expressed gene: *CLSTN2*, high expressed in human deep layer and mouse upper layer, which significantly associated with pathway of cognitive and memory behavior and reduction in functional inhibitory synapses [45], suggesting that genes, relating to cognition and memory pathways, expressed earlier in human than mouse.

Pathway analysis of two layers specific to humans revealed enrichment of axon development, such as axon guidance and axonal fasciculation, compared with mouse (Fig. 4B). Additionally, enriched pathway contained in many aspects, including visual system,