Title

Decoding the molecular landscape of the developing spatial processing system and production of entorhinal stellate cell-like cells by a direct programming approach.

Author line

Tobias Bergmann[1], Yong Liu[1,], Leo Mogus[1], Julie Lee[2,3], Ulrich Pfisterer[4], Louis-Francois Handfield[5], Andrea Asenjo-Martinez[4], Irene Lisa-Vargas[4], Stefan E Seemann[6], Jimmy Tsz Hang Lee[5], Nikolaos Patikas[5], Birgitte Rahbek Kornum[7], Mark Denham[8], Poul Hyttel[9], Menno P Witter[10], Jan Gorodkin[6], Tune H Pers[3], Martin Hemberg[5], Konstantin Khodosevich[4] and Vanessa Jane Hall*[1]

Author Affiliation

[1]Group of Brain Development and Disease, Department of Veterinary and Animal Sciences, Faculty of Health and Medical Sciences; University of Copenhagen, Frederiksberg C, 1870, Denmark.

[2]Novo Nordisk Foundation Center for Stem Cell Research, DanStem University of Copenhagen, 2200 Copenhagen, Denmark,

[3]Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences; University of Copenhagen, Copenhagen N, 2200, Denmark.

[4]Biotech Research and Innovation Centre (BRIC), Faculty of Health and Medical Sciences; University of Copenhagen, Copenhagen N, 2200, Denmark

[5]Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton CB10 1SA, United Kingdom.

[6]Center for non-coding RNA in Technology and Health, Department of Veterinary and Animal Sciences, Faculty of Health and Medical Sciences; University of Copenhagen, University of Copenhagen, 1871 Frederiksberg C, Denmark

[7]Department of Neuroscience, Faculty of Health and Medical Sciences; University of Copenhagen, Copenhagen N, 2200, Denmark.

[8]DANDRITE, Denham Group, Aarhus University, Aarhus C, 8000, Denmark

[9]Disease, Stem Cells and Embryology, Department of Veterinary and Animal Sciences, Faculty of Health and Medical Sciences; University of Copenhagen, Frederiksberg C, 1870, Denmark.

[10]Kavli Institute for Systems Neuroscience, Faculty of Medicine and Health Sciences; Norwegian University of Science and Technology, Trondheim, 7030, Norway.

Author List Footnotes:

*Corresponding Author and lead contact: Vanessa Jane Hall

Email addresses:

Yong Liu: yong.liu@ki.se
Tobias Bergmann: tobi@sund.ku.dk
Leo Mogus: rnh644@alumni.ku.dk
Julie Lee: julie.lee@sund.ku.dk
Ulrich Pfisterer: ulrich.pfisterer@med.lu.se
Louis-Francois Handfield: lh20@sanger.ac.uk
Andrea Asenjo-Martinez: andrea.asenjo@bric.ku.dk
Irene Lisa-Vargas: irene.vargas@sund.ku.dk
Stefan E Seemann: sse@sund.ku.dk
Jimmy Tsz Hang Lee: tl7@sanger.ac.uk
Nikolaos Patikas: np13@sanger.ac.uk

Birgitte Rahbek Kornum: kornum@sund.ku.dk
Mark Denham: mden@dandrite.au.dk
Poul Hyttel: poh@sund.ku.dk
Menno P Witter: menno.witter@ntnu.no
Jan Gorodkin: gorodkin@sund.ku.dk
Tune H Pers: tune.pers@sund.ku.dk
Martin Hemberg: mhemberg@bwh.harvard.edu
Konstantin Khodosevich: konstantin.khodosevich@bric.ku.dk
Vanessa Jane Hall: vh@sund.ku.dk

Corresponding Author:
Vanessa Jane Hall: vh@sund.ku.dk

Abstract 200/250 words

Classic studies investigating how and when the entorhinal cortex (component of the memory processing system of the brain) develops have been based on traditional thymidine autoradiography and histological techniques. In this study, we take advantage of modern technologies to trace at a high resolution, the cellular complexity of the developing porcine medial entorhinal cortex by using single-cell profiling. The postnatal medial entorhinal cortex comprises 4 interneuron, 3 pyramidal neuron and 2 stellate cell populations which emerge from intermediate progenitor and immature neuron populations. We discover four MGE-derived interneurons and one CGE-derived interneuron population as well as several IN progenitors. We also identify two oligodendrocyte progenitor populations and three populations of oligodendrocytes. We perform a proof-of-concept experiment demonstrating that porcine scRNA-seq data can be used to develop novel protocols for producing human entorhinal cells in-vitro. We identified six transcription factors (*RUNX1A1, SOX5, FOXP1, MEF2C, TCF3, EYA2*) important in neurodevelopment and differentiation from one *RELN*+ stellate cell population. Using a lentiviral vector approach, we reprogrammed human induced pluripotent stem cells into stellate cell-like cells which expressed *RELN, SATB2, LEF1* and BCL11B. Our findings contribute to the understanding of the formation of the brain's cognitive memory and spatial processing system and provides proof-of-concept for the production of entorhinal cells from human pluripotent stem cells in-vitro.

Keywords

Entorhinal cortex, development, stellate cell, speed cell, single-cell RNA sequencing, direct programming, induced pluripotent stem cells

Background

Uncovering the temporal events in the emergence of different cell types in the developing entorhinal cortex (EC) is paramount for understanding the formation of cognitive memory and spatial processing. The EC has a pertinent role in the processing of memory and navigation, which is dependent on its intrinsic organization and extrinsic connectivity. Specific navigation-related functions have been discovered in the EC including object recognition (1), processing speed movement (2), processing location (3), processing head-direction (4), recognizing proximal borders (5) and processing working memory (6). These functions signify the importance of the EC in spatial navigation and arise from subsets of firing cells from either the medial EC (MEC) or lateral EC (LEC). Extensive anatomical and neuronal electrophysiological classification of the EC has been performed (7-9) which has recently been supplemented with molecular characterization by single-cell RNA sequencing (scRNA-seq) in the adult EC (10, 11). Currently, none, or only very few genes can be used to identify the phenotypically identified cells, such as the grid cells, head-direction cells and speed cells and the locations of these cell types are only partially understood. In some cases, these phenotypic attributes appear to be derived from anatomically different cell types. For example, within the MEC, the grid cells are located in both the dorsal and ventral MEC (12) and found across all cortical layers, though showing a preference in layer (L) II (13). However, the grid cells are considered to represent principal neurons including both stellate cells (SCs) and pyramidal neurons (14-16) suggestive of networking interrelations that govern their function. Gene markers that distinguish the SCs from the pyramidal neurons are Reelin (*RELN*) and Calbindin (*CALB1*), respectively (17-22), although a subpopulation of both types expresses both markers (22). Reelin is also expressed in Fan cells which reside in the superficial layer of the LEC and these neurons are important for object recognition (23, 24). A specialized grid cell, the conjunctive grid cell, fires when a grid vertex is crossed upon heading in a specific direction. It is absent from LII, mostly present in LIII and LIV, but can also be found in LVI (13). The head-direction cells are principal neurons found in LII to LVI and are absent from LII (13, 25). The border cells are also considered to be principal neurons and are found within all the MEC layers (5) which often overlap with a SC identity (26). The speed cells map to GABAergic inhibitory neurons which express parvalbumin (PVALB), but not somatostatin (SST) (27).

A clearer understanding of the molecular landscape of the EC may help to reconcile the anatomical, functional and connective profiles of the cell types that exist in the EC which would further improve our understanding of spatial navigation processing. One single-cell RNA sequencing (scRNA-seq) study was recently published on the EC from Alzheimer's disease (AD) patient and healthy brains, however, the vast majority of cells captured in their dataset were oligodendrocytes (11). A second study has highlighted the transcriptional profiles of several excitatory neurons and glia in the adult EC, but does not decipher the cell types residing in the LEC or the MEC (10). Further, in both studies the research concentrated on differences between healthy and disease state rather than on the characterization of the cell types *per se*. We envisage a more comprehensive analysis of the EC cell types as they emerge during development can provide further insight into the formation processes within the EC and insight into transcriptional changes that regulate maturation and connectivity of the EC.

Studying the signaling cascades that regulate the differentiation of brain cell types has resulted in novel protocols for producing several neuronal cell subtypes in-vitro from pluripotent stem cells. Protocols exist for production of, for example, multipotent neuroepithelia (28-30), radial glia (30-38), glutamatergic neurons (39-42), GABAergic neurons (43-46), midbrain dopaminergic neurons (47-55), 5-HT neurons (56-60), motor neurons (61-64) and enteric neurons (65, 66). This process can be orchestrated via indirect differentiation using exogenous

factors and/or small molecules added to the media, or by directly programming the cells into the desired end phenotype by using targeted overexpression of regulatory transcription factors (67-69). Producing neurons of the EC from human iPSCs would provide a useful model for studying the functions of individual cell types but also be useful in modeling diseases which affect the EC, such as AD. Recent research has highlighted that both Reelin+ and *RORB*+ neurons in the EC are particularly vulnerable to either accumulating Amyloid beta, neurofibrillary tangles or dying early in the disease (10, 70) and studying these cell types further will help to elucidate why select populations of neurons in the EC are particularly vulnerable to the disease. We therefore aim to determine the key regulatory transcription factors in entorhinal Reelin+ cells in order to create a novel direct programming protocol for producing these neurons in-vitro. In light of difficulties in obtaining human fetal tissues from the second and third trimester to investigate EC development, we used an alternate large mammal, the pig. The pig develops a gyrencephalic brain midway through gestation, similar to humans and our recent research confirms it to be an excellent model of the developing human EC that recapitulates neurogenesis more closely in the human than rodents (71). We focus our attention on the MEC given the phenotypes and connectivity of several cell types are well studied but the molecular genotypes residing within the MEC are not well delineated.

In this study, we determine the molecular profiles of the neural cell types within the developing and postnatal MEC using scRNA-seq and identify entorhinal specific neuron subtypes. We then identify potential regulators of differentiation in interesting *RELN*+ excitatory neuron populations by performing differential expression of transcription factors. Focusing on one *RELN*+ population, we transduce six selected transcription factors using a lentiviral overexpression approach and directly program human induced pluripotent stem cells (iPSCs) into *RELN*+ entorhinal SC-like progenitors.

Results

Subheading 1. ScRNA-seq reveals thirty two cell populations in the developing MEC

ScRNA-seq is a powerful approach for revealing the molecular identity of cell types in the different types of tissue, including the brain. We used this methodology to investigate the timing in the emergence of different cell types of the EC and their molecular identity during development and postnatal maturation. ScRNA-seq was performed using the 10X Genomics microfluidics-based Chromium platform on whole-cells isolated from the EC at embryonal day 50 (E50), E60, E70 and postnatal day 75 (P75), and isolated nuclei in the later stages of development at E70 and P75 to ensure efficient capture of the neurons in the tissue (Figure 1A). We chose to integrate datasets for all age groups into a one-dimension reduction analysis as opposed to analyzing time points separately to ensure superior normalization and to allow us to easily cross-label cell types that existed both during fetal development and in the postnatal brain. We excluded red blood cells (expressing hemoglobin subunits *HBB, HBE1, HBM, HBZ* and on average, only 291 genes per cell) and vascular cells (*PDGFRB*$^{+}$/*PECAM1*$^{+}$) from our dataset, which left 24,294 cells (mean of 2798 different genes per cell) for further analyses. We log-normalized each sample dataset, merged all samples, batch corrected by FastMNN (72) and performed cell clustering in Seurat (Figure 1A). Clustering of the merged dataset revealed thirty-two distinct cell clusters (Figure 1B) with many clusters represented by cells from datasets across different batches with good integration of batched within same age (Figure 1C). The clusters could be delineated into 6 main cell populations demarcated by canonical markers

for intermediate progenitor cells (IPs, *PAX6*$^+$/*EOMES*$^+$/*NEUROD1*$^+$), excitatory neurons (Excitatory, *TBR1*$^+$/*EMX1*$^+$/*SATB2*$^+$/*BCL11B*$^+$/*SLC17A6*$^+$/*SLC17A7*$^+$/*DCX*$^+$), GABAergic interneurons (IN,*GAD1*$^+$/*GAD2*$^+$/*DCX*$^+$), microglia (Microglia, *AIF1*$^+$/*CXCR3*$^+$/*PTPRC*$^+$/*CSF1R*$^+$), oligodendroglia (Oligo, *MBP*$^+$/*CLDN11*$^+$/*OLIG1*$^+$/*PDGFRA*$^+$), astrocytes (*AQP4*$^+$/*GFAP*$^+$/*GLAST*$^+$) and radial glia (RGC, *HES1*$^+$/*SOX2*$^+$) (Figure 1D). Cluster 30, contained a small number of cells (n=70) which contained twice as many transcripts and genes expressed (Figure 1C) which we believed to be doublets and which failed to be removed by the data-preprocessing. Cluster 31 expressed *MBP* (Figure 1C) but was also omitted from the sub-analysis, as the highest enriched genes in this cluster were mitochondrial genes (Figure S1A) and may likely be dying cells. Furthermore, we validated our marker-expression driven annotations by projecting our dataset using scmap (73) onto a publicly available and annotated datasets of the human embryonic prefrontal cortex (74) and the human adult middle temporal gyrus (75) (Figure 1E). We found a high concordance between our annotation and the human brain datasets, despite the obvious regional differences.

Subheading 2. Progenitor and adult interneuron (IN) populations derived from the MGE and CGE are detected in the developing MEC

To investigate the temporal emergence of INs within the EC, we sub-clustered a total of 3,498 cells from the dataset using the canonical IN markers, *GAD1* and *GAD2* (clusters 2, 16, 18, 21, 22, 23 and 25) which produced ten subclusters IN0-IN9 (Figure 2A). We included clusters 2 and 25 in our IN analyses even though the scmap projection annotated these two clusters as excitatory neurons (Figure 1D, E). Despite this discrepancy, cluster 25 expressed IN genes, such as *GAD1* and *GAD2* and lacked *SLC17A6/7* and became IN7 after subclustering. For cluster 2 from the main dataset (mainly present at E60), a closer analysis of canonical markers suggested an ambiguous identity as the cells expressed both the excitatory markers, *SLC17A6/7* and IN markers, *GAD1/2* (Figure 1B, D). The UMI counts for this cluster were normal, therefore, it is unlikely to contain doublets. Cluster 2 when subclustered spread across the IN0-IN2 and IN6 populations. Five of the IN populations were only detected during development, suggesting they were IN progenitors (IN1, IN5, IN7, IN8 and IN9). IN1, IN5 and IN7 were detected during earlier embryonic gestation in the pig at E50 and E60, compared to IN8 and IN9 which were detected later, at E70 (Figure 2B). IN5, IN7 and IN8 also clustered further away from the other IN clusters which also suggests they are more transcriptionally distinct than the other IN clusters (Figure 2A). The remaining IN populations, IN0, IN2, IN3 and IN6 were all found in the postnatal brain and are likely adult IN populations (Figure 2B). We could identify several well-known IN markers that were either expressed broadly or label specific IN cell subtypes, including, *CALB1*, *DLX1*, *DLX5*, *GAD1, GAD2, LHX6*, *NPY, RELN, SST*, and *SOX6,* (Figure 2C). However, we were unable to identify other IN markers, *HTR3A, PVALB*, *CCK*, *NOS1*, *LHX5* and *VIP*, likely due to low expression levels and gene dropout effects.

Most cortical INs are generated in the ganglionic eminences and are divided in two major classes based on their place of origin: the medial ganglionic eminence (MGE) and caudal ganglionic eminence (CGE) (76). The MGE produces PVALB and SST-expressing cortical INs, which account for >60% of all INs in the cortex and the CGE produces a highly diverse group of INs, which in mice express serotonin 5HT3A receptor (77, 78) (gene *HTR3A*). However, in human INs, HTR3A labels only a minor population of CGE INs (75), which is also likely the case in pigs, since we did not detect robust *HTR3A* expression in our data. Given we could not identify PVALB in our data, we also used other MGE IN markers including *LHX6* (79) and *SOX6* (80) to identify MGE-derived INs. Based on *SST* expression in IN0, IN2 and *LHX6* and *SOX6* in IN3, we assign these clusters as MGE-derived INs (Figure 2C). It was difficult to assign IN6 to either MGE or CGE INs, given an overlapping expression profile but

previous research suggests *GRIN3A*, which was highly expressed in this population, is detected in SST INs (81), suggesting IN6 is likely of MGE origin. The progenitor population IN1 also expressed *SST*. Interestingly, both IN0 and IN1 have relatively similar transcriptional profiles. It is likely that these are unique cell populations since IN0 had lower expression of *RELN* and IN1 expressed higher levels of *RAMP3, ISG12* and *ACHE* and only IN0 could be detected after E70 of development (Figure 2C). However, further phenotypic profiling is required to conclude that these are separate populations. The remaining clusters had more distinct gene signatures. CGE-derived INs were identified based on expression of Coup TF1 (*NR2F1* gene), CoupTF2 (*NR2F2* gene) (82) and *NDNF (83)*. One population, IN4, expressed *NR2F1* and *NR2F2* and is assigned as a CGE population (Figure 2C). The remaining progenitor populations IN5, IN7-IN9 clusters also expressed these CGE IN markers (Figure 2C).

To investigate the location of the varying IN progenitors and adult populations across the EC layers, we evaluated the expression of unique genes from the clusters in the *in-situ* hybridized EC of the mouse brain from the public databases, the Allen Mouse Brain Atlas (2004) (84, 85) and the Allen Developing Mouse Brain Atlas (2008) (86). IN0 and IN1 both expressed a gene, *RESP18,* which was localized to the superficial layers and deep layers of the adult mouse EC (85) (Figure 2D), suggesting these INs are not located in the middle cortical layers. In addition *ACHE* was upregulated in IN1 and also found in the superficial and deep layers, confirming the location of at least the IN1 neurons in the mouse EC (Figure 2D). *ZIC1, ITPR1, SFRP2* and *TULP1* were expressed in the IN progenitor populations IN5, IN7, IN8 and IN9, respectively and were distributed across the layers of the developing and postnatal EC mouse brain. *ZFPM2* was upregulated in IN2 and found only in the deeper layers (Figure 2D). *MAF* and *ZBTB16* were unique upregulated genes in IN3 and IN4, respectively and were expressed across all the EC layers, however, ZBTB16 was also upregulated in the superficial layers suggesting potential high abundance of IN4 CGE INs in the superficial layers. (Figure 2D). *GRIN3A* was uniquely expressed in IN6 and detected across the layers but also upregulated in the deeper layers of the MEC suggesting different localization of these INs between the MEC and LEC (Figure 2D). In sum, we identified five IN progenitor populations (IN1, IN5, IN7-IN9), four adult IN MGE populations (IN0, IN2, IN3, IN6), and one IN CGE population (IN4).

Subheading 3. Single-cell RNA sequencing analysis reveals that OPCs emerge in the early EC.

We then assessed the oligodendroglia populations within the EC, which consisted of oligodendrocyte progenitor cells (OPCs) and oligodendrocytes at various stages of maturation. Subclustering and subsequent analyses were performed on clusters 3, 12, 24, 27 (Figure 3A insert). Five distinct clusters were identified from the subclustering (Figure 3A-C). Two clusters (OPC1 and 2) expressed OPC markers *PDGFRA* and *OLIG1* (Figure 3C), with a large proportion of the cells undergoing cell cycle division (50% and 25% respectively, Figure 3D). We were unable to identify *OLIG2* in these two clusters, likely due to low copy numbers and subsequent gene dropout effects. However, this data suggested these two clusters were progenitor cells. Cluster OPC1 was composed of mostly fetal cells from E50 to E70 whilst cluster OPC0 was detected both during gestation and postnatally (Figure 3B). OPC0 and OPC1 are clearly two separate populations, as both populations are detected after birth, albeit the clustering suggests a larger population of OPC0 after birth. Several genes have been previously identified in OPCs. For example, cluster OPC0 expressed the genes *LHFPL3* and *MMP16*, and cluster OPC1, *STMN1*; previously identified in rodent OPCs (87-90). To characterize whether these clusters resided in spatially different niches, we examined their unique gene signatures in the EC from the Allen Mouse Brain Atlas and Allen Developing Mouse Brain Atlas ISH

datasets (84-86). We detected *Cntn1* (a marker previously identified in OPCs (91)) from OPC0 in the superficial layers in the developing mouse EC (E18.5) and at a later time point across all cell layers, at P28 in the mouse EC (85, 86) (Figure 3E). We detected the OPC1 marker *Hmgb1* (also previously identified in OPC progenitors (92)) across all EC cell layers in the postnatal P4 and P28 and adult mouse EC (85, 86) (Figure 3E) showing the gene as expressed even later, after birth.

Interestingly, our data revealed an additional population, OLI2, that was constituted predominantly of fetal cells from E60 OLI2 (Figure 3A-B). These cells surprisingly also expressed the mature oligodendrocyte markers, *MBP*, *CLDN11*, *GRP17*, *NKX2-2*, and *MAG* (Figure 3C). Together, this data suggests that a small population of myelinating oligodendrocytes might be present in the fetal porcine EC already at E60. The oligodendrocyte clusters OLI3 and OLI4 consisted almost exclusively of adult cells and were highly similar in profile (Figure 3B, C). We speculate these to be mature myelinating oligodendrocytes. Nearly all the OLI3 cells expressed G2M/S-phase cycling genes while OLI4 cells were mostly quiescent and expression genes from the G0/1-phase of the cell cycle (Figure 3D). In the Allen Developing Mouse Brain Atlas, expression of the OLI2 specific marker *Fyn* showed predominant expression in the upper layers of the developing E18.5 mouse EC (86) (Figure 3E). In the Allen Mouse Brain Atlas, the OLI3 unique gene marker, *Oxr1,* was expressed in LI-III of the adult mouse EC while the OLI4 unique gene marker, *Dlg2*, was found in the deep layers (85) (Figure 3E-F). In summary, our data indicate the presence of 2 OPC populations that may populate different layers of the EC, with one pre-myelinating oligodendrocyte population and two mature oligodendrocyte profiles with similar profiles.

Subheading 4. Single-cell RNA sequencing reveals unique gene signatures for stellate cells and pyramidal neuron populations

To gain a deeper understanding of the temporal pattern of emerging neurons we performed subclustering on the IP and excitatory neuron populations identified from the parent dataset and uncovered unique gene expression signatures. We subclustered and analyzed a total of 6,833 cells, including the excitatory clusters 0, 5, 10, 15, and 17 and the IP cluster 19. Figure 4A, insert). The subclustering revealed ten distinct populations (Figure 4A). Cells from the gestational stages dominated the subset of IPs and excitatory neurons, with only very few adult cells represented (Figure 4B). This is due to the fact that we captured mostly glia in the postnatal brain samples. Amongst the clusters, one population exhibited a typical IP identity. Cells in cluster IP4 expressed the typical IP markers *EOMES, SOX2*, and *NEUROD4* (Figure 4C) (93) and could be detected across all stages of development, but also in the postnatal brain suggestive of ongoing postnatal neurogenesis (Figure 4B). IP8 was a detected only during the neurogenesis period (E50-E60) (71) and expressed typical neural progenitor markers including, *PAX3* (94), *DBX1* (95) and the ventral midbrain progenitor marker *OTX2* (96). When referring to the Allen Developing Mouse Brain Atlas ISH dataset (86), the E18.5 mouse EC showed expression of a unique marker *Cxcr4*, across all the layers (Figure S2). *Pax3* at P4 and *Dbx1* at P14 are observed expressed in the superficial layers (Figure S2) suggesting that the IP8 population consists of superficial neuron progenitors.

An analysis of canonical genes for immature neurons with a pyramidal-like identity revealed five clusters. Clusters PYR0, PYR1, PYR2, PYR3 and PYR5 were tangentially located on the t-SNE plot (Figure 4A) and expressed low levels of *RELN* (Figure 4C). Surprisingly, we did not observe in our dataset any neuron populations expressing *CALB1*, a common pyramidal neuron marker reported in rats, mice and humans within the EC (97-99). *CALB1* may not be a

suitable marker for pyramidal neurons in the pig EC as a previous study has reported the absence of expression in pyramidal neurons in the hippocampal region (100). Clusters PYR0 and PYR1 shared similar transcriptional profiles but diverged in *NEUROD1* expression. Clusters PYR2 and PYR3 were also similar, but had divergent *ARPP21* expression. Only PYR1 and PYR3 could be detected in the prenatal brain and not in the postnatal brain. Given the similar profiles, it is likely that PYR0, PYR2 and PYR5 are adult populations. This does not explain however why all five populations are identified during early development, so we cannot eliminate the possibility that these are potentially different pyramidal neurons. Therefore, we deduce to have captured at least three pyramidal neuron populations. An analysis of mature pyramidal neuron markers only partially confirmed the notion of maturation of PYR1, PYR3 and PYR5, as *SLC17A6/7* (vGLUT1/2 genes) expression was observed in PYR0 and PYR2 but not PYR5 (Figure 4C). Validation of two unique gene markers in the Allen Developing Brain Mouse Atlas (86) corroborated the putative location of these cells. The gene *TNR* is highly expressed in PYR2 and PYR3 and was found in the superficial layers of the mouse EC (Figure S2). The gene *NEUROD6* was highly upregulated in PYR2, PYR3 and PYR5 and can be detected in both the superficial and deeper layers. Our investigations of the pig EC show BCL11B is expressed both in LII and LV of the EC (71)) and we found *BCL11B* to be highly upregulated in PYR5 suggesting that cluster PYR5 may reside in both the superficial and deep layers. Taken together, we have identified five pyramidal neuron subtypes during development and at least three mature pyramidal neuron types in the postnatal EC that are uniquely placed across the EC layers.

In addition to the pyramidal neurons, the EC harbors two types of principal neurons expressing Reelin. These are the stellate and fan cells that reside in LII/III in the MEC and LEC, respectively (22). Our subclustering analysis of the MEC revealed three distinct neuron populations that expressed *RELN* (RELN 6, RELN7 and RELN9), two of which were present postnatally (RELN6 and RELN7) (Figure 4 A,B). Apart from the known expression of Reelin and absence of CALB1 (at least in a proportion of Reelin neurons), few other genes are known to be expressed in these cells (21, 101, 102). RELN9 was only present during gestation (mostly at E60-E70) (Figure 4B) and clustered closely to PYR0 and IP4. It expressed *DLX1*, *DLX5* and *ARX* (Figure 4A-C), together with *GAD1/2* (Figure S1B), suggesting it has more of a GABAergic IN progenitor (103-105)-like identity and we deduce this not to be a reelin excitatory neuron.

This left us with clusters RELN6 and RELN7, which lacked the pyramidal transcription factor, *EMX1*(106) (Figure 4C). These two clusters were expressed throughout all sampled time points and persisted in the postnatal adult brain (Figure 4B). They had distinct transcriptional profiles and were distinctly separated on the t-SNE plot (Figure 4A, B). Interestingly, RELN6 cells are more abundant than RELN7 cells at E50 suggesting that more of this RELN subtype is born during early neurogenesis (Figure 4B). In contrast, more RELN7 cells were captured at E70 (late neurogenesis in the pig (71)) than RELN6 cells (Figure 4B). When assessing unique markers from these clusters in the Allen mouse atlases, the RELN6 uniquely identified gene, *Lhx9*, was expressed across all the EC layers but also across the parahippocampal cortex at E18.5 and P14 in the mouse (86) (Figure S2). This marker is known to be expressed in Reelin negative (-) pioneer neurons across the cortex (107) and is therefore not an ideal marker for identifying the RELN6 SC within the MEC alone. In the Allen Developing Mouse Brain Atlas, another marker, *Ndst4* was expressed at higher levels in the superficial layers at P56 in the mouse EC (86) (Figure 4D) which indicates their location within the superficial layers. When assessing genes from cluster RELN7 in the Allen Developing Mouse Brain Atlas, moderate

levels of *Reln* and *Cpne4*, which are both expressed in the superficial layers of the mouse EC and were more enriched in the LEC than the MEC LII (86) (Figure 4C, D). Further, *Epha6* was upregulated in cluster RELN7 and could be observed in the LII-LIII in the Allen Mouse Brain Atlas (85), confirming the RELN7 neurons are also located in the superficial layer. To sum, two *RELN* excitatory neuron populations were identified which persisted after birth.

Subheading 5. Induced expression of six transcription factors in human iPSCs results in stellate cell-like progenitors

The Reelin+ SCs are a vulnerable cell population in the EC (10, 70) and would be an interesting cell type to model in-vitro for studying AD. We sought to use our scRNA-seq data to produce a novel protocol for the production of SCs from pluripotent stem cells by induced overexpression of SC specific transcription factors. Therefore, looked for candidate TFs among the twenty most enriched TFs in the two *RELN* populations based on the TFcheckpoint list of DNA-binding RNA polymerase II TFs curated by Chawla *et. al* (108) using the Findmarkers tool in Seurat (Figure 5A). We focused our attention on the RELN7 cluster, given it represented a mature SC profile that was located in the superficial layers of the EC. *RORB* was found to be one of the twenty differentially expressed TFs, and it has recently been identified as a marker for vulnerable EC excitatory neurons (10). Of the 20 TFs we focused on six, *RUNX1A1, SOX5, FOXP1, MEF2C, TCF4* and *EYA2* based on a literature search showing prominent roles for these TFs in neurodevelopment and differentiation. Runx1t1 is specifically expressed in neurons and plays a role in hippocampal neuron differentiation (109, 110), which also lies in close proximity with the EC within the ventral telencephalon and both structures arise from the medial pallium (111). Eya2 is expressed in neural progenitors and is repressed by Foxg1 during corticogenesis, playing an important role in neuronal differentiation (112). Foxp1 is co-expressed with Satb2 and detected in LIII-LVa neurons (113). It is also implicated in neural stem cell differentiation and neuronal morphogenesis (114, 115). We recently demonstrated that MEC SC express SATB2 which further raised our interest in FOXP1 (71). *SATB2* was also detected in both RELN7 and RELN9 (Figure 4C). SOX5 is an important regulator of early-born deep layer neurons (116) and in differentiation of specific corticofugal neuron subtypes (117) but may be involved in SC fate given they also express the deep layer marker, BCL11B (71). *BCL11B* was also upregulated in the RELN7 population and more so than RELN6 (Figure 4C). Tcf4 is expressed in cortical and hippocampal neurons, highly enriched in the pallial region (118) and is involved in neurogenesis, neuronal differentiation and hippocampal formation (119-121). It is also implicated in memory and spatial learning (122). Mef2c plays a role in neuronal development and is expressed in hippocampal neurons (123). It is also implicated in cognitive learning, including working memory and object recognition (124). An assessment of these genes in the mouse brain ISH data from the Allen Mouse Brain Atlas demonstrated that both *Tcf4* and *Mef2c* are expressed in the superficial layers of the EC, whereas the remaining genes were expressed across all the layers (85) (Figure 5B). The sequences of these six TFs were also available from Addgene and cloned into plasmids from either mouse (Foxp1, Sox5) or human (TCF4, MEF2C, RUNX1T1, EYA2) RNA.

In order to produce SC from human iPSCs, we applied an overexpression approach in combination with culture conditions that induced a medial pallial fate. Evidence suggests the MEC is derived from the medial pallium (111, 125, 126). We also tested whether Ngn2 was necessary in the direct programming approach since Chen *et al.* have demonstrated that patterning neurons with specific factors in combination with ectopic Ngn2 expression yielded

a more uniform population with a specific CNS fate (127). We performed lentiviral transduction in the human iPSC lines, SFC 180-01-01 (SFC) and WTSli024A (WTS) using different combinations of lentiviruses expressing the TFs in the presence or absence of Ngn2 (applying a leave-one-out approach in the presence of Ngn2) and cultured the transduced cells in CHIR and BMP4 for 5 days (Figure 5C). Puromycin selection was initiated on day (D)2 and continued 5 days to eliminate non-transduced cells. The cells were re-seeded on D5 and propagated in neuronal medium for up to a further 15 days (total 20 days) (Figure 5C). We also included a negative control where the human iPSCs were not transduced with TFs. We determined that these cells were unable to differentiate into a neuronal fate in the medial pallial culture conditions alone (Figure 6). As to be expected, the Ngn2 alone treatment resulted in neuronal morphology at D10 and D20, as has been described previously (128). The combination of the six SC TFs in the presence of Ngn2 also resulted in neurons with axonal projections at D20 (Figure 6A). However, this was only observed in the SFC line, and pronounced cell death was observed in the WTS cell line, which might be attributed to cell line-specific toxicity. Strikingly, the combination of the six SC TFs in the absence of Ngn2 resulted in bipolar cells at D10, and at D20 a more neuron-like morphology was observed (Figure 6A). The leave-one-out SC TF approach in the presence of Ngn2 also resulted in neurons with extending dendrites at D20 (Figure S3A). By morphological observation alone, we were not able to exclude the contribution of any particular, or combinations of SC TF in directing a neuronal fate in the presence of Ngn2.

In order to further elucidate the identity of the neurons produced, we investigated the expression of known SC markers, *BCL11B, SATB2, RELN,* and genes upregulated in the MEC LII, *LEF1* and *LHX2* (71, 111, 129, 130) in the cells reprogrammed with all six SC TFs with or without Ngn2, compared to Ngn2 alone, at D10 and D20. We selected *BCL11B* since we have determined earlier that most earlyborn SC are BCL11B+ (71), however other SC populations are BCL11B-. The qPCR confirmed expression of *BCL11B* in all three treatments, but the highest levels of expression were observed in the Ngn2-only induced cells (Figure 6B). A significant increase in *SATB2* expression was also observed in all three treatments, but was highest in the Ngn2-only neurons (Figure 6B). Interestingly, a significant increase ($p < 0.001$) in *RELN* expression was observed only in the cells induced with expression of the SC TFs alone (Figure 6B). Together, the SC TFs induced co-expression of *BCL11B, SATB2*, and *RELN*, mirroring the unique molecular profile of EC SCs. To rule out a pyramidal neuron identity we analyzed the cells for *CALB1*. Of interest, was a high expression of the pyramidal marker, *CALB1* in the cells induced with the SC TFs alone. This was unexpected, as only a smaller subpopulation of entorhinal SCs, the intermediate SCs, express both Reelin and Calbindin (22). Further, *LEF1* was upregulated in both treatments containing SC TFs but not in the Ngn2 alone treatment (Figure 6B). In contrast, *LHX2* was only upregulated in the Ngn2-only treatment. The analyses indicate the SC TFs were able to induce the expression of *BCL11B, SATB2*, *RELN, CALB1* and *LEF1* in the human iPSCs derived neurons by D20. To further determine a SC glutamatergic identity, we assessed the iPSC-derived cell populations at D10 and D20 for expression of BCL11B, Reelin, and vGLUT1 by performing immunocytochemistry. Interestingly, none of the neurons from the different treatments expressed Reelin or vGLUT1 (data not shown). The absence of Reelin was not surprising since expression of this gene is switched on later in development. However, BCL11B expression was observed in cells transduced with Ngn2, SC TFs and Ngn2 and SC TFs alone, at D20 (Figure 6C) and as early as D10 in the WTS line (Figure S3B). This observation indicates that the SC TFs alone were able to induce BCL11B expression in cells with neuronal morphology. Omission of one of the six SC TFs in the presence of Ngn2 also resulted in BCL11B expression in the cells (Figure S4). To sum, the SC TFs alone were able to differentiate human iPSCs into SC-like cells with neuronal processes and several unique SC markers. Further, *RELN* was only detectable at the

RNA level which suggests an immature neuronal phenotype given the lack of vGLUT1 expression.

Discussion

Our study demonstrates the cellular architecture of the developing EC and provides a proof-of-concept approach for using scRNA-seq data to produce region-specific neurons from induced pluripotent stem cells. We demonstrate that scRNA-seq data from the developing and postnatal brain can be used to identify transcriptional regulators important for specification of neuronal cell subtypes. We identified in a proof-of-concept study that six TFs: *FOXP1, SOX5, TCF4, MEF2C, RUNX1T1* and *EYA2*, when added in combination, resulted in the differentiation of entorhinal SC-like cells. Further studies are required to identify if one or more of these TFs may be redundant in the reprogramming process. This could have been facilitated in hindsight if Ngn2 had been removed from the leave-one-out experiment. Nevertheless, we prove that transcriptional regulators of neuronal subtype cell fate can be identified from scRNA-seq clusters which are defined in the curated TFcheckpoint list and extracted using the Findmarkers tool in Seurat. This study will fuel future studies for streamlining the reprogramming process and optimizing the transduction process.

The phenotype of the SC-like cells produced may reflect the intermediate SC which is an entorhinal cell type that expresses both Reelin and Calb (21, 22). Interestingly, we could not detect *CALB1* in the RELN7 cluster, which could be attributed to low gene expression and threshold cut-offs of low expressing genes. Our own and other studies have shown peculiarities of CALB expression in the pig ventral telencephalon. For example, CALB1 protein is localized only in white matter tracts in the EC of the postnatal pig EC and could not be detected at all, prior to E100 (two weeks prior to birth) (71). Further inherent differences show a lack of CALB1 in mossy fibers, granule cells and pyramidal neurons in the pig hippocampus (100). Post transcriptional mechanisms might be responsible for a lack of protein expression in pyramidal neurons in the EC, such as upstream repressors of the open reading frame (131), frameshifting events (132) or alternative transcription start sites that create different transcript isoforms (133). The upregulation of *RELN* expression in the SC-like cells points to a SC phenotype as it is rarely expressed in neurons. An exception to the entorhinal SC are the glutamatergic granule cells in the cerebellum (134). Instead *RELN* is predominantly expressed in GABAergic interneurons (134), but the combination of genes in these cells did not reflect an interneuron phenotype. Instead, the co-expression of BCL11B and SATB2 highly favors a superficial or deep layer entorhinal neuron phenotype in the MEC (71), however LEF1 is highly enriched in the MEC LII (135) which points the cells towards having a superficial layer identity. The lack of Reelin protein expression is not surprising since we see protein expression of Reelin turns on several days (approx. 10 days in the pig) after birth of the SC (71), which may even be longer in the human, given the longer gestation period. Prolonged cultures of these iPSC-derived neurons and more extensive marker evaluation could help to confirm their identity. We cannot rule out that the RELN6 population are also stellate cells residing in the superficial layers. These cells also expressed *BCL11B* but were *RORB* negative and expressed only low levels of *SATB2*. Our previous studies in the mouse EC show Reelin+/Bcl11b+ neurons are born in the superficial layers (LII and LIII) from E12.5 to E15, but most were born at E12.5 (71). Another subset of Reln+/Bcl11b- were also born from E12.5 to E15 (71). In the pig scRNA-seq data, both RELN6 and RELN7 expressed similar low levels of *BCL11B* but appear to be born at different time points when assessing the varying proportions of the cells captured at E50 and E70. This suggests RELN6 cells may arise earlier than RELN7 cells. However, we did not capture a *RELN*+ population expressing high levels of *BCL11B* in the postnatal brain

samples so it is difficult to confer our previous mouse data to our current porcine scRNA-seq data. Further studies analyzing differences between the RELN populations may help to distinguish the RELN subtypes better.

Our study also demonstrates that pig scRNA-seq data can be readily applied for generation of novel protocols in human iPSCs. Indeed we found the developing EC dataset projected extremely well using scmap to human and mouse fetal datasets. This was the case, even when the human datasets were at unmatched gestational ages and from the embryonic prefrontal cortex and adult middle temporal gyrus. Our previous findings have also identified that the pig poses as an excellent model of neurogenesis, given its long gestational length, completion of corticogenesis prior to birth and presence of a moderate outer subventricular zone (71).

Our single-cell profiling of the porcine postnatal EC revealed a high quality dataset which corroborates and further enhances knowledge of EC cell types in the postnatal brain. Another scRNA-seq study was recently published on the human adult EC from both AD patient and healthy brains, however, their data showed a median of 646 genes captured in 13,214 sequenced cells (11) Similarly, Leng et al., profiled a higher number of cells in the human caudal EC (42,528 cells) however the depth of sequencing resulted in a variable number of genes captured per cell, for e.g. a mean of 2,000 genes per excitatory cell) (10). Our dataset resulted in a mean of 2,798 different genes captured per cell in the 24,294 cells. In a report by Grubman et al., the vast majority of cells captured in their dataset were predominantly glia with very few neurons (11). For example in their study, only six neuronal subtypes could be identified in the adult EC. The Leng et al. study extended this knowledge and identified nine excitatory neuron populations. Our study identified twelve excitatory neurons, some of which exist only during development, but furthers the field's knowledge on the EC subtypes. We also characterized the neurons based on the two main neuron types of the EC, the pyramidal and SC subtypes. We identified at least three subtypes of pyramidal neurons and two stellate neuron subtypes in the postnatal MEC. We also identified four IN subtypes in the postnatal MEC which is fewer than the nine inhibitory neuron populations identified in Leng et al., in the adult human brain (10) but were able to detect six IN progenitor populations, when also including the *RELN*+ population that clustered closely to the excitatory neurons in the excitatory neuron subclustering analyses. The variation in numbers of IN populations seen in our study compared to the Leng study may be due to species-specific differences, sampling differences or differences in clustering methodology. In regions where IN diversity have been well studied such as the amygdala, at least 6 classes of INs have been identified based on their discharge properties and electrophysiological profiles in the basolateral amygdala (136) and similar research is required in the EC to fully differentiate between the different IN subclasses. The INs lying in the superficial layers are particularly interesting for researchers studying the mechanisms underlying grid cell firing. We found expression of unique genes from IN4 and IN0 to be highly abundant in the superficial layers of the mouse EC, but of the two populations, only IN4 is SST- which suggests this may be the IN population of most interest, since SST-/PVALB+ INs are directly coupled to grid cell firing (137). We also confer with the Leng et al. study, that two oligodendrocyte populations exist in the postnatal brain (10). The two previous scRNA-seq studies on the EC have concentrated on differences between healthy and diseased brains and characterization of vulnerable cell populations to AD rather than on the characterization of the cell types *per se*. Our dataset therefore provides excellent insight into the cell diversity of the EC due to the increased cell capture rate, depth of sequencing, and subclustering analyses.

In conclusion, we demonstrate that scRNA-seq data from the developing brain can be used to produce novel protocols for direct programming of pluripotent stem cells into subtype specific

entorhinal cell fates and we provide evidence in a proof-of-concept approach to produce SC-like cells. Production of novel cell types ex-vivo will facilitate research into tissues, such as the EC, which is afflicted early on in e.g. AD. We envisage future studies could further build upon this dataset by adding the spatial dimension and neuron connectivity to build even more accurate and detailed maps of the spatial processing system.

METHODS

Aim, design and setting of the study

Animal welfare and collection of brains

The experiments conform to the relevant regulatory standards for use of fetal material. Crossbred male and female pig fetuses at Embryonic day (E)50, E60, and E70 of development were obtained from inseminated Danish Landrace/Yorkshire sows with Duroc boar semen from a local pig production farm. Deceased postnatal pigs were obtained at postnatal day (P)75 as a gift from Per Torp Sangild at the University of Copenhagen. Adult brains were obtained from sows killed for another study using an overdose of sodium phenobarbital by a professional issued with a license from the Danish Animal Experiment Inspectorate.

Single-cell preparation

In total we prepared 10 sequencing-libraries form isolated MECs from E50 (whole cell, 3 brains, 1 batch), E60 (whole cell, 4 brains, 3 batches), E70 (whole cell, 4 brains, 3 batches; nucleus 1 brain, one batch) and from adult sow MEC (whole cell, 1 brain, 1 batch; nucleus, 1 brain, 1 batch) (Table S1). Briefly, the individual MEC tissue was digested using a papain dissociation method, according to the manufacturer's guidelines (Worthington) with small modifications.

The EC was macroscopically dissected out (approx. 1 mm^3) in the digest medium (1x PBS (Thermo Fisher Scientific), 1x Penicillin-Streptomycin (Sigma-Aldrich)), transferred to a 3.5cm Petri dish and incubated in 1mL papain solution for 30 min at 37^{o}C. The tissue was gently triturated 20 times. The cell suspension was diluted with 1 mL FBS (BioWest) and centrifuged for 5 min at 300 g at room temperature (RT). The supernatant was discarded and the cell pellet was resuspended in 2.7 mL digestion media (1x Neurobasal medium (Thermo Fisher Scientific), 10% FBS (BioWest), 1x Penicillin-Streptomycin (Sigma-Aldrich)), 300 uL albumin-ovomucoid inhibitor, and 150 uL DNAse solution. The cell suspension was carefully layered on top of 5 ml of albumin-inhibitor solution in a 15 mL falcon tube and centrifuged for 6 min at 70 g at RT. The supernatant was discarded and the cell pellet was resuspended in 5 mL cell-resuspension media (1x Neurobasal medium (Thermo Fisher Scientific), B27 (Thermo Fisher Scientific), 1x Penicillin-Streptomycin (Sigma-Aldrich), bFGF (5 ng/ml, Prospec). The cells were counted (NucleoCounter, ChemoMetec) and ranged in viability from 80.4 - 88.4% (Viability and Cell Count assay), diluted to 100-2000 cells/uL used for single-cell library preparation.

Single-nuclei isolation

Nuclei extraction was performed as described before (138) with the following modifications (139). Prior to nuclei extraction, nuclei isolation medium 1 (NIM1) (250 mM sucrose, 25 mM KCl, 5 mM MgCl2, 10 mM Tris Buffer pH8), NIM2 (NIM1 buffer supplemented with 1 µM DTT (Thermo Fisher Scientific) and 1x EDTA-free protease inhibitors (Roche) and homogenization (NIM2 buffer supplemented with Recombinant RNase Inhibitor (0.4 U/µL, Takara), SUPERase in (0.2 U/µL, Thermo Fisher Scientific) and Triton (0.1% v/v, Sigma-Aldrich)) buffers were prepared. Briefly, sectioned frozen brain tissue was placed into pre-cooled 1ml dounce homogenizer (Wheaton) with 1ml ice-cooled homogenization buffer. Tissue was dissociated on ice using 5-6 strokes with the loose pestle and 15-17 strokes with the tight pestle. Homogenate was first filtered through a 70 µm filter. Nuclei were collected (900 g, 10 min) and resuspended in 500µl staining buffer (nuclease free PBS (1X, Thermo Fisher Scientific), BSA (0.5% wt/vol, Sigma-Aldrich), SUPERase in (0.2 U/µL, Thermo Fisher Scientific)). The sample was stained with 7-AAD (2ug/uL, Sigma-Aldrich) in order to visualize nuclei during FACS sorting. 7-AAD positive nuclei were FACS-isolated (70 µm nozzle, BD Biosciences, BD FACSAria™) into a 1.5ml eppendorf tube containing 10 µl 10% nuclease free BSA (Thermo Fisher Scientific). Single 7-AAD+ nuclei were isolated using the gating strategy similar to Pfisterer et al (139).

Single-nuclei RNA-seq library preparation and sequencing

Whole cells were loaded onto the 10X Genomics microfluidic chip according to the Chromium Single Cell 3' Reagent Kits User Guide version 2 chemistry (10X Genomics). The single-nuclei samples were loaded as per whole cells, albeit the samples were not diluted. 10-12.000 thousand cells/nuclei were loaded from each sample.

Libraries from two samples at different stages were pooled and sequenced together on an Illumina NextSeq 500 (Table S1) following the NextSeq System Denature and Dilute Libraries Guide Protocol A: Standard Normalization Method (illumina). The NextSeq 500/550 High Output Reagent Cartridge v2 75 cycles (illumina) kit was used for the whole cell and single-nuclei samples and the pooled library were sequenced on NextSeq 500. The sequencing cycles were: Read1, 26 cycles, i7 Index 8 cycles, i5 Index 0 cycles; Read2, 57 cycles.

Plasmid design and construction

The selected genes of interest (GOIs) were cloned into the pTet-O-Ngn2-puro plasmid (Addgene #52047) by replacing *Ngn2* with the GOI. This ensured that the GOI was fused to the puromycin resistance gene. Plasmids containing *EYA2* (Addgene #49264), RUNX1T1 (Addgene #49264), *TCF4* (Addgene #109144), *Foxp1* (Addgene #16362), *Sox5* (Addgene #48707) and *MEF2C* (Addgene #61538) were all acquired from Addgene. Forward and reverse primers were designed for subcloning all GOIs (with the exception of MEF2C which was already subcloned into a lentiviral Tet-O plasmid (Addgene #61538) (see Table S2 for primer sequences). Gene inserts were amplified from the plasmids by PCR amplification using 400 - 600 ng template plasmid, 0.5 µM forward and reverse primer (Eurofins Genomics), 5 µL Pfu DNA Polymerase 10X Buffer (Promega), 1 µL Pfu DNA Polymerase (Promega), and 200 µM dNTP mix (Promega) in a 50 µL reaction volume. The following program was used in a PTC-200 Thermal Cycler (MJ Research): 1 cycle of 2 min at 95 °C, 40 cycles of 1 min at 95 °C, 30 sec at 60 °C, and 1 min at 72 °C. The PCR amplified inserts were run on a 1% agarose gel and purified using the Wizard SV Gel and PCR Clean-Up System kit (Promega) according to the manufacturer's instructions. The gene insert was then ligated into a Zero Blunt TOPO vector using the Zero Blunt TOPO PCR Cloning Kit (Invitrogen) according to the manufacturer's protocol. The TOPO vector containing the gene insert was heat-shock transformed into

competent E. coli (NEB Stable C3040I), and grown at 37 °C O/N on agar (Sigma-Aldrich) plates containing 50 mg/mL kanamycin (Sigma-Aldrich) for selection. Individual colonies were propagated at 37 °C O/N in 5 mL Luria-Bertani (LB) broth (Sigma-Aldrich) supplemented with 50 mg/mL kanamycin. Plasmids were isolated using the PureYield Plasmid Miniprep System (Promega) according to the manufacturer's protocol. The GOIs were isolated from the purified plasmids by digestion with EcoRI (Esp3I for RUNX1T1) and XbaI according to the manufacturer's protocol (FastDigest, Thermo Fisher Scientific). Positive clones were purified from a 1% agarose gel. The GOI was then inserted and ligated into a dephosphorylated pTet-O-Ngn2-puro (addgene 52047) backbone to create the final plasmid; pTet-O-GOI-puro using the LigaFast Rapid DNA Ligation System (Promega) and heat-shock transformed into competent E. coli. Individual colonies were selected and isolated using the PureYield Plasmid Miniprep System (Promega) according to the manufacturer's protocol. Correct plasmid inserts were validated by digestion with EcoRI and XbaI (FastDigest, Thermo Scientific). The plasmids were verified by Sanger sequencing (Eurofins Genomics) and stored at −80 °C until use for virus production.

Lentiviral production and titering

Second generation lentiviruses were produced following transfection of HEK293FT cells. On the day of transfection, HEK293FT cells were cultured in fresh DMEM supplemented with 10% newborn bovine calf serum (NCS) (Hyclone, New Zealand) in the absence of antibiotics. The cells were transfected as follows: In one tube, 50 µL Lipofectamine 3000 reagent (Invitrogen) was diluted in 1.5 mL Opti-MEM (Gibco; Thermo Fisher Scientific). 40 µL P3000 reagent (Invitrogen) and 5 pmol DNA plasmids (1.5 pmol psPAX2 (Addgene #12260), 1.5 pmol pMD2.G (Addgene #12259) and 2 pmol transfer plasmid containing the GOI) were diluted in 1.5 mL Opti-MEM. The two tubes were mixed and incubated at RT for 10 min and the solution was then added dropwise to the cells. After 24 h the media was replaced with fresh DMEM supplemented with 10% NCS and incubated for 48 h. The virus was harvested 72 h post transfection. The virus was spun at 500 g for 5 min and the supernatant filtered through a 45 µm polyethersulfone filter (Fisher Scientific) and subsequently concentrated, by spinning at 23,000 rpm (using rotor Beckman Type 60 Ti) for 2 h at 4 °C. The supernatant was discarded and 100 µL PBS was added to the pellet and incubated for 1 h at 4 °C. The pellet was resuspended by pipetting and stored at −80 °C until use for transduction.

Titering of virus was performed on HEK293FT cells by adding lentivirus at different amounts (from 1 to 15 µL virus). A virus with a predetermined titer was added as a reference and a negative control without virus was included. On Day 1, the media was substituted with fresh media, and 48 h (Day 2) after transduction, the genomic DNA (gDNA) was extracted from the cells using the PureLink Genomic DNA Mini Kit (Invitrogen) according to the manufacturer's protocol. Each sample was analyzed in triplicate by qPCR using both primers for a sequence fused to the GOI (LV2), fused to all GOIs, and for the β-actin intron (see Table S3 for primer sequences). The LV2 was used to determine the degree of viral integration into the genome. The β-actin gene was used to determine the number of genome copies, as an estimate for the total number of cells. For the qPCR reaction, 5 µL SYBR green (Roche), 1 µL 10 mM of both forward and reverse primer and 15 ng gDNA was mixed in a 10 µL reaction volume and loaded in a white qPCR plate with optical caps (Thermo Fisher Scientific). Samples were spun briefly and run on a Mx3005P qPCR machine (Agilent) using the following program: 1 cycle of 10 min at 95 °C, 40 cycles of 20 sec at 95 °C, 20 sec at 55 °C, and 30 sec at 72 °C, followed by a melting curve analysis.

The viral titers were determined based on a recently published approach with minor variations (140). Briefly Ct values were converted from their intrinsic exponential relation to linear related Xo values, using the Xo method (141). The triplicate LV2 X o mean values for a sample was normalized to its mean Xo value for β-actin to obtain the copy number of integrated GOI relative to β-actin gene copies in the given sample. A standard curve of normalized LV2 sequence copy number versus the volume of virus that was added to the sample initially was fitted for each titration of the virus, respectively. All curves followed a linear relationship with R2 ranging from 0.889 to 0.995. The relative GOI integration for the reference virus was then calculated and inserted into the linear regression curves for each respective virus, where (VPD, volume of produced virus): VPD = a * relative copy number of reference virus + b. The VPD is a measure of the volume of the produced virus that is needed, under the same conditions as the reference virus, to obtain the same degree of GOI integration as was obtained for the reference virus. The VPD for each of the produced viruses was hence calculated by inserting the relative copy number of the reference virus in the linear regression for the respective virus produced. The titer of the produced virus was finally calculated based on the titer of the reference virus, using a multiplicity of infection (MOI) of 10, where U is units of viral particles: VPD = (Number of cells * MOI) / Titer (U / µL).

Cell culture

For direct reprogramming, two human iPSC lines were selected including the human SFC180-01-01 (female donor, healthy, age unknown) and WTSli024A (female donor, healthy, age 50-54) (both acquired from the EBiSC cell bank). The cells were maintained in Geltrex (Gibco; Thermo Fisher Scientific) coated dishes with mTESR media (Stemcell Technologies) supplemented with 1% Penicillin-Streptomycin (Gibco; Thermo Fisher Scientific). The cells were passaged using 0.5 mM EDTA (Merck) and propagated until they reached 50-60% confluency, before initiating reprogramming using the produced lentivirus. The cell lines were cultured at 37 °C in 5% CO2 with 90% humidity.

Transduction of human iPSCs

Human iPSCs were transduced in mTESR in the absence of antibiotics and with 6 µg/mL polybrene (Millipore). The lentiviruses containing the GOIs were transduced using 10 MOI in the presence of the reverse tetracycline transactivator M2rtTA (MOI 10). After 3 h the media was supplemented with half the usual volume of fresh media. The day after viral transduction was considered Day 0, and the media was exchanged for medial pallium (MP) patterning media (47% Neurobasal A medium, 47% DMEM/F12+Glutamax, 1% Insulin-Transferrin-Selenium A (Gibco; Thermo Fisher Scientific), 2% B-27 Supplement (Gibco; Thermo Fisher Scientific), 1% CTS N-2 supplement (Gibco; Thermo Fisher Scientific), 0.5% GlutaMAX (Thermo Fisher Scientific), 1% Pen Strep, 0.3% Glucose, 3 µM CHIR 99021 (Sigma-Aldrich), 5 nM BMP4 (R&D systems), 1 µg/mL doxycycline (Sigma-Aldrich)). On Day 2 and 3 the media was exchanged for MP patterning media supplemented with 3 µg/mL puromycin to select for transduced cells. On Day 5, the cells were dissociated using dispase (Sigma-Aldrich), spun for 5 min at 300 g and cultured in NBM media (96% Neurobasal medium, 2% B-27 Supplement, 1% GlutaMAX, 1% pen strep, 1 µg/mL doxycycline, 200 µM Ascorbic Acid (Sigma-Aldrich), 1 µM DAPT (Sigma-Aldrich)) and 3 µg/mL puromycin. Upon confluency, the cells were split onto triple coated (1x polyornithine (Sigma) 10 µg/mL fibronectin (Sigma-Aldrich),10 µg/mL PolyD laminin (Sigma-Aldrich) plates. The cells were propagated in the NBM medium up until D20.

Immunocytochemistry

The differentiated human iPSCs were fixed on glass coverslips (VWR, Denmark) in 4% PFA

for 15 min at RT and stored at 4ºC until use. The cells were washed in PBS and permeabilized in 0.25% Triton X-100 in PBS, washed once in PBS-T (0.1% Tween-20 (Sigma-Aldrich) in PBS) and then treated with blocking buffer (10% NDS, 3% BSA, 0.1% Tween-20 in PBS) for 30 min at RT. Immunocytochemistry was performed using a combination of primary antibodies targeted against BCL11B (1:400, Abcam, ab18465), Reelin (1:100, Santa-cruz, sc-25346) and vGLUT1 (1:200, Abcam, ab79774) diluted in blocking buffer 2 for 60 min at RT. The cells were washed 3x 5 min in PBS-T and incubated with secondary antibodies conjugated with Alexa fluorophores (1:1000, Invitrogen, A10036, A21208, and A21448) diluted in blocking buffer for 60 min at RT. The cells were then washed twice for 5 min in PBS-T and counterstained for 10 min with Hoechst 33342 (1 µg/mL in PBS). The cells were washed 3x 5 min in PBS, dipped 3x in deionized water and mounted onto microscope slides using buffered glycerol mounting media 90% glycerol (Sigma-Aldrich), 20mM Tris pH 8.0, 0.5% N-propyl gallate (Sigma-Aldrich).

Image processing and quantification

For acquisition of immunocytochemistry images, a confocal microscope Leica TCS SPE was used. Images were optimized for brightness and contrast using Fiji-ImageJ. Statistical analysis was conducted using the commercially available software Prism 7.0 (GraphPad Software).

qPCR

Total RNA was extracted from cells using TRIzol reagent (Life Technologies) according to the manufacturer's protocol. After isolation of the aqueous phase 10 µg of glycogen (Thermo Fisher Scientific) was supplemented before precipitation of the RNA. The pellet was resuspended in nuclease free water and the concentration was determined using a NanoDrop. Synthesis of cDNA was performed using the GoScript Reverse Transcriptase kit (Promega) with random hexamer primers, according to the manufacturer's protocol. Lastly, samples were heat inactivated for 10 min at 70 °C and diluted in nuclease free water. The qPCR was performed in biological triplicates. For the qPCR reaction, 5 µL SYBR green (Roche), 1 µL 10 mM of both forward and reverse primer (see Table S3 for primer sequences) and 10 µL cDNA was mixed in a 10 µL reaction volume and loaded in a white qPCR plate with optical caps (Thermo Fisher Scientific). Samples were spun down briefly and run on a Mx3005P qPCR machine (Agilent) using the following program: 1 cycle of 10 min at 95 °C, 40 cycles of 20 sec at 95 °C, 20 sec at 52 °C, and 30 sec at 72 °C, followed by a melting curve analysis. The Ct values of all samples were first converted from their intrinsic exponential relation to linear related Xo values, using the Xo method (141). In brief, the triplicate LV2 X o mean values for a sample was normalized to its mean Xo value for GAPDH. The SD was calculated accounting for the common coefficient of variation both for the housekeeping gene and the individual gene investigated. Values represent the gene expression normalized to the expression of GAPDH for each individual sample + SD. Statistical one-way ANOVA,with Tukey's multiple comparisons test was performed in GraphPad Prism, comparing the expression for one gene between all different samples.

QUANTIFICATION AND STATISTICAL ANALYSIS

Initial quality control and data analysis

Briefly, the sequenced libraries were mapped to pre-mRNA and filtered using 10X Genomics Cell Ranger pipeline. The sequencing data was demultiplexed by bcl2fast (Illumina) which warped in the Cell Ranger followed by aligning to reference genome (Sscrofa11.1 release-94) by STAR (142). Finally, mRNA molecules were counted by Cell Ranger. The quality of the

sequencing libraries was assessed by Cell Ranger, which determined in each sample the sequencing depth cutoff that is required for cells to be included in the downstream analysis. All samples were merged for downstream clustering and data analysis, which was performed in Seurat v2 (143). Similar numbers of UMIs and genes were observed across all batches (Figure S5A, B). Overview of number of cells, mean reads per cell, median genes per cell, and number of reads and pooled brains in the included scRNA-seq libraries can be found in Table S4.

Batch correction

We used decomposeVar from the 'scran' R package (RRID: SCR_001905) in order to find a list of variable genes that are used for PCA dimensionality reduction; however, to aid the correspondence of single-cell to single-nucleus for down-stream analysis (such as clustering), we identified and excluded genes whose transcripts were highly abundant in empty droplets Empty droplets were defined as cells with <50 UMIs. If the number transcripts that are found in the empty droplet is above 30% the total number of transcripts for a given gene (Figure S5C), that gene will be removed from the list of variable genes.

The individual datasets were log-normalized and the PCA projected gene expression in the datasets were then batch corrected by FastMNN (72), which reduced the distance between cell pairs that are found to be reciprocally nearest neighbors across batches (before batch correction; Figure S5D, after correction; Figure 1C). The merged dataset were visualized by t-distributed stochastic neighbour embedding (t-SNE) (144).

Unsupervised clustering and cell type identification

Cell types were identified using the Louvain algorithm (145) with a resolution parameter set at 1.6 for the entire dataset, 1.4 for the IN analysis, 1.2 for the oligodendrocyte analysis, and 1.0 for the IP and excitatory neuron analysis . The canonical markers were used to identify the neurons of the clusters (Figure 1D). In addition, we used two reference datasets which each contained smart-seq2 single-cells from human embryonic prefrontal cortex (74, 146) or from human adult middle temporal gyrus (147). We used scmap v1.12.0 (73) to project individual cells onto curated cell-type clusters that are available in each reference. Each cell-type prediction utilizes the consensus of three similarity measures from queried cell to reference cluster centroids using sets of cell-type markers that were identified in the respective reference datasets; however, only the human genes that possess an ortholog gene in pig were used as cell type marker for similarity measure calculation.

Differential gene analysis

Marker genes were identified using scfind v3.7.0 (PMID 33649586) based on highest F1 score for cluster specific genes. In principle by: For cluster cj, gene gi is considered a true positive (TP) if it is expressed, a false negative (FN) if it is not expressed, a false positive (FP) if is expressed in a cell assigned to another cluster, and a true negative (TN) if it is not expressed in a cell assigned to a different cluster. For each gi we evaluate precision = TP/(TP+FP), recall = TP/(TP+FN), and F1 = 2*precision*recall/(precision + recall). For each cj, genes are ranked by F1 with the highest scoring genes used as markers.

Identification of cell type specific transcription factors

Four excitatory neuron clusters expressing *RELN*, determined as potential SC populations, were investigated for cluster-specific enriched TFs. Porcine genes were defined as TFs (total of 2475 genes) if they, or orthologues, were present in the TFcheckpoint, a list of 3479 specific DNA-binding RNA polymerase II TFs curated by Chawla and colleagues (108). Differentially expressed TFs between each *RELN* cluster and all remaining cells were identified using

FindMarkers in the Seurat pipeline (148), which uses Wilcoxon Rank Sum test. For each of the 2 *RELN* clusters, 20 TFs were selected with the highest log fold-change of average expression between the cluster and the remaining cells.

Cell cycle score analysis

Cell cycle phases were scored using the built-in function in Seurat CellCycleScoring, which was used to determine whether a given cell was likely to be in either S, G2M or G1 (which is indistinguishable from G0) phase of cell cycle. Cell cycle scores were based on a list of cell cycle phase-specific genes proposed by Tirosh *et al*. (149).

Code availability

All R code used for the analysis of this study is publically available at github: https://github.com/BDD-lab/EC-scRNAseq-2021

Abbreviations

BMP – bone morphogenetic protein
CHIR - GSK-3α/β inhibitor
CNS – central nervous system
D – Day
DMEM – dulbecco's modified eagle's medium
E – Embryonic day
EC – entorhinal cortex
EDTA - Ethylenediaminetetraacetic acid
gDNA – genomic DNA
GOI – gene of interest
h - hours
iPSC –induced pluripotent stem cells
IN – interneurons
IP – intermediate progenitor
L – layer
LV - lentivirus
LEC – lateral entorhinal cortex
MEC – medial entorhinal cortex
MOI – multiplicity of infection
OPC – oligodendrocyte progenitor cell
P – postnatal day
PBS – phosphate-buffered saline
PCR – polymerase chain reaction
PYR – pyramidal neuron
qPCR – quantitative real time PCR
RELN – Reelin gene
RT – room temperature
scRNA-seq – Single-cell RNA sequencing
SC – stellate cell
SD – standard deviation
SPE – speed cell
SST – somatostatin

TF – transcription factor
U – units of viral particles
UMI – unique molecular identifier
VPD – volume of produced virus

Declarations:

Ethics approval and consent to participate

The experiments conform to the relevant regulatory standards for acquisition and use of fetal pig material.

Consent for publication

All authors consent that this research is suitable for publication.

Availability of data and materials

The datasets generated are available at the NCBI repository with GEO accession number GSE134482 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE134482). The submitted data includes the raw sequencing data as fastq files together with the processed count matrix used in this study.

Competing interests

The authors declare that they have no competing interests.

Funding

The project was financed by The Independent Research Fund, Denmark under the grant (ID: DFF–7017-00071 and 8021-00048B), Lundbeck fund (R296-2018-2287) and The Innovation Foundation (Brainstem (4108-00008B)). Konstantin Khodosevich acknowledges the Novo Nordisk Hallas-Møller Investigator grant (NNF16OC0019920). Tune Pers acknowledges the Novo Nordisk Foundation (Grant number NNF18CC0034900) and the Lundbeck Foundation (Grant number R190-2014-3904). Martin Hemberg and Nikolaos Patikis were supported by a core grant to the Wellcome Sanger Institute from the Wellcome Trust. Louis-Francois Handfield was supported by Open Targets (OTAR038) and Jimmy Lee was supported by a CZI (2018-183503 (5022)) from the Chan Zuckerberg Initiative.

Authors' contributions

TB: Performed scRNA-seq, direct programming, analyzed data and contributed to the writing of the manuscript
YL: Performed scRNA-seq, analyzed data, designed the GOIs subclone pipeline and contributed to the writing of the manuscript
LM: Performed lentiviral production.
JL: Performed scRNA-seq analysis.
UP: Performed scRNA-seq analysis
L-FH: Contributed to methodology of bioinformatic analysis of the scRNA-seq data
AAM: Performed scRNA-seq analysis
ILV; Performed scRNA-seq analysis
MD: Supplied Ngn2 backbones and Ngn2 lentivirus
SES: Contributed to methodology of bioinformatic analysis of the scRNA-seq data
JTHL: Contributed to methodology of bioinformatic analysis of the scRNA-seq data
NP: Contributed to methodology of bioinformatic analysis of the scRNA-seq data

MP: Contributed to methodology of bioinformatic analysis of the scRNA-seq data
BRK: Contributed to methodlogy of scRNA-seq
MD: Contributed with lentiviral plasmids for direct programming experiments and viral tittering methodology
PH: Supervised project and contributed to acquisition of funding
MPW: Contributed to anatomical analysis of the EC in the pig and in acquisition of funding
JG: Provided server access for scRNA-seq data, supervised the project and contributed to interpretation of bioinformatic data
MH: Provided software, personnel and resources for bioinformatic analyses of the scRNA-seq and contributed to interpretation of the scRNA-seq data
THP: Provided materials, personnel and resources for scRNA-seq, supervised project and contributed to interpretation of the scRNA-seq data
KK: Provided materials, personnel and resources for scRNA-seq, supervised project and contributed to interpretation of the data
VH: Contributed to the project's concept, methodology, formal analyses, project's resources, writing, reviewing & editing the manuscript. Administered and managed the project and acquired the project's funding.
All authors read and approved the final manuscript

Acknowledgements
We thank Per Torp Sangild for his donation of healthy adult sow brains for this study. We used the Single Cell Genomics Facility at BRIC for access to 10X Chromium platform. We also acknowledge the help from Anna Fossum at the FACS Facility at BRIC, University of Copenhagen for help with the FACS sorting.

Declaration of Interests:
The authors declare no competing interests.

References

1. Ridley RM, Samson NA, Baker HF, Johnson JA. Visuospatial learning impairment following lesion of the cholinergic projection to the hippocampus. Brain Res. 1988;456(1):71-87.
2. Kropff E, Carmichael JE, Moser MB, Moser EI. Speed cells in the medial entorhinal cortex. Nature. 2015;523(7561):419-24.
3. Quirk GJ, Muller RU, Kubie JL, Ranck JB, Jr. The positional firing properties of medial entorhinal neurons: description and comparison with hippocampal place cells. J Neurosci. 1992;12(5):1945-63.
4. Taube JS, Muller RU, Ranck JB, Jr. Head-direction cells recorded from the postsubiculum in freely moving rats. II. Effects of environmental manipulations. J Neurosci. 1990;10(2):436-47.
5. Solstad T, Boccara CN, Kropff E, Moser MB, Moser EI. Representation of geometric borders in the entorhinal cortex. Science. 2008;322(5909):1865-8.
6. Olton DS, Walker JA, Wolf WA. A disconnection analysis of hippocampal function. Brain Res. 1982;233(2):241-53.
7. Canto CB, Witter MP. Cellular properties of principal neurons in the rat entorhinal cortex. II. The medial entorhinal cortex. Hippocampus. 2012;22(6):1277-99.
8. Canto CB, Witter MP. Cellular properties of principal neurons in the rat entorhinal cortex. I. The lateral entorhinal cortex. Hippocampus. 2012;22(6):1256-76.

9. Canto CB, Wouterlood FG, Witter MP. What does the anatomical organization of the entorhinal cortex tell us? Neural Plast. 2008;2008:381243.
10. Leng K, Li E, Eser R, Piergies A, Sit R, Tan M, et al. Molecular characterization of selectively vulnerable neurons in Alzheimer's disease. Nat Neurosci. 2021.
11. Grubman A, Chew G, Ouyang JF, Sun G, Choo XY, McLean C, et al. A single-cell atlas of entorhinal cortex from individuals with Alzheimer's disease reveals cell-type-specific gene expression regulation. Nat Neurosci. 2019;22(12):2087-97.
12. Stensola T, Moser EI. Grid Cells and Spatial Maps in Entorhinal Cortex and Hippocampus. In: Buzsaki G, Christen Y, editors. Micro-, Meso- and Macro-Dynamics of the Brain. Cham (CH)2016. p. 59-80.
13. Sargolini F, Fyhn M, Hafting T, McNaughton BL, Witter MP, Moser MB, et al. Conjunctive representation of position, direction, and velocity in entorhinal cortex. Science. 2006;312(5774):758-62.
14. Domnisoru C, Kinkhabwala AA, Tank DW. Membrane potential dynamics of grid cells. Nature. 2013;495(7440):199-204.
15. Rowland DC, Obenhaus HA, Skytoen ER, Zhang Q, Kentros CG, Moser EI, et al. Functional properties of stellate cells in medial entorhinal cortex layer II. Elife. 2018;7:e36664.
16. Schmidt-Hieber C, Hausser M. Cellular mechanisms of spatial navigation in the medial entorhinal cortex. Nat Neurosci. 2013;16(3):325-31.
17. Perez-Garcia CG, Gonzalez-Delgado FJ, Suarez-Sola ML, Castro-Fuentes R, Martin-Trujillo JM, Ferres-Torres R, et al. Reelin-immunoreactive neurons in the adult vertebrate pallium. J Chem Neuroanat. 2001;21(1):41-51.
18. Seress L, Leranth C, Frotscher M. Distribution of calbindin D28k immunoreactive cells and fibers in the monkey hippocampus, subicular complex and entorhinal cortex. A light and electron microscopic study. J Hirnforsch. 1994;35(4):473-86.
19. Varga C, Lee SY, Soltesz I. Target-selective GABAergic control of entorhinal cortex output. Nat Neurosci. 2010;13(7):822-4.
20. Leitner FC, Melzer S, Lutcke H, Pinna R, Seeburg PH, Helmchen F, et al. Spatially segregated feedforward and feedback neurons support differential odor processing in the lateral entorhinal cortex. Nat Neurosci. 2016;19(7):935-44.
21. Fuchs EC, Neitz A, Pinna R, Melzer S, Caputi A, Monyer H. Local and Distant Input Controlling Excitation in Layer II of the Medial Entorhinal Cortex. Neuron. 2016;89(1):194-208.
22. Witter MP, Doan TP, Jacobsen B, Nilssen ES, Ohara S. Architecture of the Entorhinal Cortex A Review of Entorhinal Anatomy in Rodents with Some Comparative Notes. Front Syst Neurosci. 2017;11:46.
23. Germroth P, Schwerdtfeger WK, Buhl EH. Morphology of identified entorhinal neurons projecting to the hippocampus. A light microscopical study combining retrograde tracing and intracellular injection. Neuroscience. 1989;30(3):683-91.
24. Nilssen ES, Jacobsen B, Fjeld G, Nair RR, Blankvoort S, Kentros C, et al. Inhibitory Connectivity Dominates the Fan Cell Network in Layer II of Lateral Entorhinal Cortex. J Neurosci. 2018;38(45):9712-27.
25. Giocomo LM, Stensola T, Bonnevie T, Van Cauter T, Moser MB, Moser EI. Topography of head direction cells in medial entorhinal cortex. Curr Biol. 2014;24(3):252-62.
26. Tang Q, Burgalossi A, Ebbesen CL, Ray S, Naumann R, Schmidt H, et al. Pyramidal and stellate cell specificity of grid and border representations in layer 2 of medial entorhinal cortex. Neuron. 2014;84(6):1191-7.

27. Ye J, Witter MP, Moser MB, Moser EI. Entorhinal fast-spiking speed cells project to the hippocampus. Proc Natl Acad Sci U S A. 2018;115(7):E1627-E36.
28. Li W, Sun W, Zhang Y, Wei W, Ambasudhan R, Xia P, et al. Rapid induction and long-term self-renewal of primitive neural precursors from human embryonic stem cells by small molecule inhibitors. Proc Natl Acad Sci U S A. 2011;108(20):8299-304.
29. Falk A, Koch P, Kesavan J, Takashima Y, Ladewig J, Alexander M, et al. Capture of neuroepithelial-like stem cells from pluripotent stem cells provides a versatile system for in vitro production of human neurons. PLoS One. 2012;7(1):e29597.
30. Nat R, Nilbratt M, Narkilahti S, Winblad B, Hovatta O, Nordberg A. Neurogenic neuroepithelial and radial glial cells generated from six human embryonic stem cell lines in serum-free suspension and adherent cultures. Glia. 2007;55(4):385-99.
31. Gorris R, Fischer J, Erwes KL, Kesavan J, Peterson DA, Alexander M, et al. Pluripotent stem cell-derived radial glia-like cells as stable intermediate for efficient generation of human oligodendrocytes. Glia. 2015;63(12):2152-67.
32. Duan L, Peng CY, Pan L, Kessler JA. Human pluripotent stem cell-derived radial glia recapitulate developmental events and provide real-time access to cortical neurons and astrocytes. Stem Cells Transl Med. 2015;4(5):437-47.
33. Liour SS, Kraemer SA, Dinkins MB, Su CY, Yanagisawa M, Yu RK. Further characterization of embryonic stem cell-derived radial glial cells. Glia. 2006;53(1):43-56.
34. Onorati M, Camnasio S, Binetti M, Jung CB, Moretti A, Cattaneo E. Neuropotent self-renewing neural stem (NS) cells derived from mouse induced pluripotent stem (iPS) cells. Mol Cell Neurosci. 2010;43(3):287-95.
35. Bamba Y, Shofuda T, Kanematsu D, Nonaka M, Yamasaki M, Okano H, et al. Differentiation, polarization, and migration of human induced pluripotent stem cell-derived neural progenitor cells co-cultured with a human glial cell line with radial glial-like characteristics. Biochem Biophys Res Commun. 2014;447(4):683-8.
36. Reinchisi G, Limaye PV, Singh MB, Antic SD, Zecevic N. Neurogenic potential of hESC-derived human radial glia is amplified by human fetal cells. Stem Cell Res. 2013;11(1):587-600.
37. Hofrichter M, Nimtz L, Tigges J, Kabiri Y, Schroter F, Royer-Pokora B, et al. Comparative performance analysis of human iPSC-derived and primary neural progenitor cells (NPC) grown as neurospheres in vitro. Stem Cell Res. 2017;25:72-82.
38. Bouhon IA, Joannides A, Kato H, Chandran S, Allen ND. Embryonic stem cell-derived neural progenitors display temporal restriction to neural patterning. Stem Cells. 2006;24(8):1908-13.
39. Cao SY, Hu Y, Chen C, Yuan F, Xu M, Li Q, et al. Enhanced derivation of human pluripotent stem cell-derived cortical glutamatergic neurons by a small molecule. Sci Rep. 2017;7(1):3282.
40. Reyes JH, O'Shea KS, Wys NL, Velkey JM, Prieskorn DM, Wesolowski K, et al. Glutamatergic neuronal differentiation of mouse embryonic stem cells after transient expression of neurogenin 1 and treatment with BDNF and GDNF: in vitro and in vivo studies. J Neurosci. 2008;28(48):12622-31.
41. Thoma EC, Wischmeyer E, Offen N, Maurus K, Siren AL, Schartl M, et al. Ectopic expression of neurogenin 2 alone is sufficient to induce differentiation of embryonic stem cells into mature neurons. PLoS One. 2012;7(6):e38651.
42. Chuang JH, Tung LC, Lee-Chen GJ, Yin Y, Lin Y. An approach for differentiating uniform glutamatergic neurons from mouse embryonic stem cells. Anal Biochem. 2011;410(1):149-51.

43. Yang N, Chanda S, Marro S, Ng YH, Janas JA, Haag D, et al. Generation of pure GABAergic neurons by transcription factor programming. Nat Methods. 2017;14(6):621-8.
44. Sun AX, Yuan Q, Tan S, Xiao Y, Wang D, Khoo AT, et al. Direct Induction and Functional Maturation of Forebrain GABAergic Neurons from Human Pluripotent Stem Cells. Cell Rep. 2016;16(7):1942-53.
45. Barretto N, Zhang H, Powell SK, Fernando MB, Zhang S, Flaherty EK, et al. ASCL1- and DLX2-induced GABAergic neurons from hiPSC-derived NPCs. J Neurosci Methods. 2020;334:108548.
46. Shin E, Palmer MJ, Li M, Fricker RA. GABAergic neurons from mouse embryonic stem cells possess functional properties of striatal neurons in vitro, and develop into striatal neurons in vivo in a mouse model of Huntington's disease. Stem Cell Rev Rep. 2012;8(2):513-31.
47. Kikuchi T, Morizane A, Doi D, Magotani H, Onoe H, Hayashi T, et al. Human iPS cell-derived dopaminergic neurons function in a primate Parkinson's disease model. Nature. 2017;548(7669):592-6.
48. Jo J, Xiao Y, Sun AX, Cukuroglu E, Tran HD, Goke J, et al. Midbrain-like Organoids from Human Pluripotent Stem Cells Contain Functional Dopaminergic and Neuromelanin-Producing Neurons. Cell Stem Cell. 2016;19(2):248-57.
49. Daadi MM. Differentiation of Neural Stem Cells Derived from Induced Pluripotent Stem Cells into Dopaminergic Neurons. Methods Mol Biol. 2019;1919:89-96.
50. Jovanovic VM, Salti A, Tilleman H, Zega K, Jukic MM, Zou H, et al. BMP/SMAD Pathway Promotes Neurogenesis of Midbrain Dopaminergic Neurons In Vivo and in Human Induced Pluripotent and Neural Stem Cells. J Neurosci. 2018;38(7):1662-76.
51. Kriks S, Shim JW, Piao J, Ganat YM, Wakeman DR, Xie Z, et al. Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. Nature. 2011;480(7378):547-51.
52. Xue Y, Zhan X, Sun S, Karuppagounder SS, Xia S, Dawson VL, et al. Synthetic mRNAs Drive Highly Efficient iPS Cell Differentiation to Dopaminergic Neurons. Stem Cells Transl Med. 2019;8(2):112-23.
53. Tofoli FA, Semeano ATS, Oliveira-Giacomelli A, Goncalves MCB, Ferrari MFR, Veiga Pereira L, et al. Midbrain Dopaminergic Neurons Differentiated from Human-Induced Pluripotent Stem Cells. Methods Mol Biol. 2019;1919:97-118.
54. Zhang P, Xia N, Reijo Pera RA. Directed dopaminergic neuron differentiation from human pluripotent stem cells. J Vis Exp. 2014(91):51737.
55. Studer L. Derivation of dopaminergic neurons from pluripotent stem cells. Prog Brain Res. 2012;200:243-63.
56. Lu J, Zhong X, Liu H, Hao L, Huang CT, Sherafat MA, et al. Generation of serotonin neurons from human pluripotent stem cells. Nat Biotechnol. 2016;34(1):89-94.
57. Kumar M, Kaushalya SK, Gressens P, Maiti S, Mani S. Optimized derivation and functional characterization of 5-HT neurons from human embryonic stem cells. Stem Cells Dev. 2009;18(4):615-27.
58. Vadodaria KC, Marchetto MC, Mertens J, Gage FH. Generating human serotonergic neurons in vitro: Methodological advances. Bioessays. 2016;38(11):1123-9.
59. Shimada T, Takai Y, Shinohara K, Yamasaki A, Tominaga-Yoshino K, Ogura A, et al. A simplified method to generate serotonergic neurons from mouse embryonic stem and induced pluripotent stem cells. J Neurochem. 2012;122(1):81-93.

60. Bethea CL, Reddy AP, Pedersen D, Tokuyama Y. Expression profile of differentiating serotonin neurons derived from rhesus embryonic stem cells and comparison to adult serotonin neurons. Gene Expr Patterns. 2009;9(2):94-108.
61. Sances S, Bruijn LI, Chandran S, Eggan K, Ho R, Klim JR, et al. Modeling ALS with motor neurons derived from human induced pluripotent stem cells. Nat Neurosci. 2016;19(4):542-53.
62. Bianchi F, Malboubi M, Li Y, George JH, Jerusalem A, Szele F, et al. Rapid and efficient differentiation of functional motor neurons from human iPSC for neural injury modelling. Stem Cell Res. 2018;32:126-34.
63. Dimos JT, Rodolfa KT, Niakan KK, Weisenthal LM, Mitsumoto H, Chung W, et al. Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. Science. 2008;321(5893):1218-21.
64. Chipman PH, Toma JS, Rafuse VF. Generation of motor neurons from pluripotent stem cells. Prog Brain Res. 2012;201:313-31.
65. Barber K, Studer L, Fattahi F. Derivation of enteric neuron lineages from human pluripotent stem cells. Nat Protoc. 2019;14(4):1261-79.
66. Workman MJ, Mahe MM, Trisno S, Poling HM, Watson CL, Sundaram N, et al. Engineered human pluripotent-stem-cell-derived intestinal tissues with a functional enteric nervous system. Nat Med. 2017;23(1):49-59.
67. Oh Y, Jang J. Directed Differentiation of Pluripotent Stem Cells by Transcription Factors. Mol Cells. 2019;42(3):200-9.
68. Li K, Kong Y, Zhang M, Xie F, Liu P, Xu S. Differentiation of pluripotent stem cells for regenerative medicine. Biochem Biophys Res Commun. 2016;471(1):1-4.
69. Qin H, Zhao A, Fu X. Small molecules for reprogramming and transdifferentiation. Cell Mol Life Sci. 2017;74(19):3553-75.
70. Kobro-Flatmoen A, Nagelhus A, Witter MP. Reelin-immunoreactive neurons in entorhinal cortex layer II selectively express intracellular amyloid in early Alzheimer's disease. Neurobiol Dis. 2016;93:172-83.
71. Liu Y, Bergmann T, Mori Y, Peralvo Vidal JM, Pihl M, Vasistha NA, et al. Development of the entorhinal cortex occurs via parallel lamination during neurogenesis. 2021.
72. Haghverdi L, Lun ATL, Morgan MD, Marioni JC. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. Nat Biotechnol. 2018;36(5):421-7.
73. Kiselev VY, Yiu A, Hemberg M. scmap: projection of single-cell RNA-seq data across data sets. Nat Methods. 2018;15(5):359-62.
74. Zhong S, Zhang S, Fan X, Wu Q, Yan L, Dong J, et al. A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex. Nature. 2018;555(7697):524-8.
75. Hodge RD, Bakken TE, Miller JA, Smith KA, Barkan ER, Graybuck LT, et al. Conserved cell types with divergent features in human versus mouse cortex. Nature. 2019;573(7772):61-8.
76. Gelman DM, Marin O. Generation of interneuron diversity in the mouse cerebral cortex. Eur J Neurosci. 2010;31(12):2136-41.
77. Rudy B, Fishell G, Lee S, Hjerling-Leffler J. Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons. Dev Neurobiol. 2011;71(1):45-61.
78. Vucurovic K, Gallopin T, Ferezou I, Rancillac A, Chameau P, van Hooft JA, et al. Serotonin 3A receptor subtype as an early and protracted marker of cortical interneuron subpopulations. Cereb Cortex. 2010;20(10):2333-47.

79. Liodis P, Denaxa M, Grigoriou M, Akufo-Addo C, Yanagawa Y, Pachnis V. Lhx6 activity is required for the normal migration and specification of cortical interneuron subtypes. J Neurosci. 2007;27(12):3078-89.
80. Batista-Brito R, Rossignol E, Hjerling-Leffler J, Denaxa M, Wegner M, Lefebvre V, et al. The cell-intrinsic requirement of Sox6 for cortical interneuron development. Neuron. 2009;63(4):466-81.
81. Murillo A, Navarro AI, Puelles E, Zhang Y, Petros TJ, Perez-Otano I. Temporal Dynamics and Neuronal Specificity of Grin3a Expression in the Mouse Forebrain. Cereb Cortex. 2021;31(4):1914-26.
82. Kanatani S, Yozu M, Tabata H, Nakajima K. COUP-TFII is preferentially expressed in the caudal ganglionic eminence and is involved in the caudal migratory stream. J Neurosci. 2008;28(50):13582-91.
83. Tasic B, Yao Z, Graybuck LT, Smith KA, Nguyen TN, Bertagnolli D, et al. Shared and distinct transcriptomic cell types across neocortical areas. Nature. 2018;563(7729):72-8.
84. Lein ES, Hawrylycz MJ, Ao N, Ayres M, Bensinger A, Bernard A, et al. Genome-wide atlas of gene expression in the adult mouse brain. Nature. 2007;445(7124):168-76.
85. Allen Mouse Brain Atlas: Allen Institute for Brain Science; 2004 [Available from: https://mouse.brain-map.org/.
86. Allen Developing Mouse Brain Atlas: Allen Institute for Brain Science; 2008 [Available from: https://developingmouse.brain-map.org/.
87. Artegiani B, Lyubimova A, Muraro M, van Es JH, van Oudenaarden A, Clevers H. A Single-Cell RNA Sequencing Study Reveals Cellular and Molecular Dynamics of the Hippocampal Neurogenic Niche. Cell Rep. 2017;21(11):3271-84.
88. Hu JG, Wang YX, Zhou JS, Chen CJ, Wang FC, Li XW, et al. Differential gene expression in oligodendrocyte progenitor cells, oligodendrocytes and type II astrocytes. Tohoku J Exp Med. 2011;223(3):161-76.
89. Lin G, Mela A, Guilfoyle EM, Goldman JE. Neonatal and adult O4(+) oligodendrocyte lineage cells display different growth factor responses and different gene expression patterns. J Neurosci Res. 2009;87(15):3390-402.
90. Magri L, Gacias M, Wu M, Swiss VA, Janssen WG, Casaccia P. c-Myc-dependent transcriptional regulation of cell cycle and nucleosomal histones during oligodendrocyte differentiation. Neuroscience. 2014;276:72-86.
91. Lamprianou S, Chatzopoulou E, Thomas JL, Bouyain S, Harroch S. A complex between contactin-1 and the protein tyrosine phosphatase PTPRZ controls the development of oligodendrocyte precursor cells. Proc Natl Acad Sci U S A. 2011;108(42):17498-503.
92. Nicaise AM, Wagstaff LJ, Willis CM, Paisie C, Chandok H, Robson P, et al. Cellular senescence in progenitor cells contributes to diminished remyelination potential in progressive multiple sclerosis. Proc Natl Acad Sci U S A. 2019;116(18):9030-9.
93. Pollen AA, Nowakowski TJ, Chen J, Retallack H, Sandoval-Espinosa C, Nicholas CR, et al. Molecular identity of human outer radial glia during cortical development. Cell. 2015;163(1):55-67.
94. Blake JA, Ziman MR. Pax genes: regulators of lineage specification and progenitor cell maintenance. Development. 2014;141(4):737-51.
95. Lacin H, Zhu Y, Wilson BA, Skeath JB. dbx mediates neuronal specification and differentiation through cross-repressive, lineage-specific interactions with eve and hb9. Development. 2009;136(19):3257-66.

96. Puelles E, Annino A, Tuorto F, Usiello A, Acampora D, Czerny T, et al. Otx2 regulates the extent, identity and fate of neuronal progenitor domains in the ventral midbrain. Development. 2004;131(9):2037-48.
97. Beall MJ, Lewis DA. Heterogeneity of layer II neurons in human entorhinal cortex. J Comp Neurol. 1992;321(2):241-66.
98. Diekmann S, Nitsch R, Ohm TG. The organotypic entorhinal-hippocampal complex slice culture of adolescent rats. A model to study transcellular changes in a circuit particularly vulnerable in neurodegenerative disorders. J Neural Transm Suppl. 1994;44:61-71.
99. Fujimaru Y, Kosaka T. The distribution of two calcium binding proteins, calbindin D-28K and parvalbumin, in the entorhinal cortex of the adult mouse. Neurosci Res. 1996;24(4):329-43.
100. Holm IE, Geneser FA, Zimmer J, Baimbridge KG. Immunocytochemical demonstration of the calcium-binding proteins calbindin-D 28k and parvalbumin in the subiculum, hippocampus and dentate area of the domestic pig. Prog Brain Res. 1990;83:85-97.
101. Kitamura T, Pignatelli M, Suh J, Kohara K, Yoshiki A, Abe K, et al. Island cells control temporal association memory. Science. 2014;343(6173):896-901.
102. Winterer J, Maier N, Wozny C, Beed P, Breustedt J, Evangelista R, et al. Excitatory Microcircuits within Superficial Layers of the Medial Entorhinal Cortex. Cell Rep. 2017;19(6):1110-6.
103. Friocourt G, Parnavelas JG. Identification of Arx targets unveils new candidates for controlling cortical interneuron migration and differentiation. Front Cell Neurosci. 2011;5:28.
104. Wang Y, Dye CA, Sohal V, Long JE, Estrada RC, Roztocil T, et al. Dlx5 and Dlx6 regulate the development of parvalbumin-expressing cortical interneurons. J Neurosci. 2010;30(15):5334-45.
105. Pla R, Stanco A, Howard MA, Rubin AN, Vogt D, Mortimer N, et al. Dlx1 and Dlx2 Promote Interneuron GABA Synthesis, Synaptogenesis, and Dendritogenesis. Cereb Cortex. 2018;28(11):3797-815.
106. Chan CH, Godinho LN, Thomaidou D, Tan SS, Gulisano M, Parnavelas JG. Emx1 is a marker for pyramidal neurons of the cerebral cortex. Cereb Cortex. 2001;11(12):1191-8.
107. Bertuzzi S, Porter FD, Pitts A, Kumar M, Agulnick A, Wassif C, et al. Characterization of Lhx9, a novel LIM/homeobox gene expressed by the pioneer neurons in the mouse cerebral cortex. Mech Dev. 1999;81(1-2):193-8.
108. Chawla K, Tripathi S, Thommesen L, Laegreid A, Kuiper M. TFcheckpoint: a curated compendium of specific DNA-binding RNA polymerase II transcription factors. Bioinformatics. 2013;29(19):2519-20.
109. Zou L, Li H, Han X, Qin J, Song G. Runx1t1 promotes the neuronal differentiation in rat hippocampus. Stem Cell Res Ther. 2020;11(1):160.
110. Linqing Z, Guohua J, Haoming L, Xuelei T, Jianbing Q, Meiling T. Runx1t1 regulates the neuronal differentiation of radial glial cells from the rat hippocampus. Stem Cells Transl Med. 2015;4(1):110-6.
111. Abellan A, Desfilis E, Medina L. Combinatorial expression of Lef1, Lhx2, Lhx5, Lhx9, Lmo3, Lmo4, and Prox1 helps to identify comparable subdivisions in the developing hippocampal formation of mouse and chicken. Front Neuroanat. 2014;8:59.
112. Kumamoto T, Toma K, Gunadi, McKenna WL, Kasukawa T, Katzman S, et al. Foxg1 coordinates the switch from nonradially to radially migrating glutamatergic subtypes in the neocortex through spatiotemporal repression. Cell Rep. 2013;3(3):931-45.

113. Hisaoka T, Nakamura Y, Senba E, Morikawa Y. The forkhead transcription factors, Foxp1 and Foxp2, identify different subpopulations of projection neurons in the mouse cerebral cortex. Neuroscience. 2010;166(2):551-63.
114. Braccioli L, Vervoort SJ, Adolfs Y, Heijnen CJ, Basak O, Pasterkamp RJ, et al. FOXP1 Promotes Embryonic Neural Stem Cell Differentiation by Repressing Jagged1 Expression. Stem Cell Reports. 2017;9(5):1530-45.
115. Li X, Xiao J, Frohlich H, Tu X, Li L, Xu Y, et al. Foxp1 regulates cortical radial migration and neuronal morphogenesis in developing cerebral cortex. PLoS One. 2015;10(5):e0127671.
116. Kwan KY, Lam MM, Krsnik Z, Kawasawa YI, Lefebvre V, Sestan N. SOX5 postmitotically regulates migration, postmigratory differentiation, and projections of subplate and deep-layer neocortical neurons. Proc Natl Acad Sci U S A. 2008;105(41):16021-6.
117. Lai T, Jabaudon D, Molyneaux BJ, Azim E, Arlotta P, Menezes JR, et al. SOX5 controls the sequential generation of distinct corticofugal neuron subtypes. Neuron. 2008;57(2):232-47.
118. Kim H, Berens NC, Ochandarena NE, Philpot BD. Region and Cell Type Distribution of TCF4 in the Postnatal Mouse Brain. Front Neuroanat. 2020;14:42.
119. Mesman S, Bakker R, Smidt MP. Tcf4 is required for correct brain development during embryogenesis. Mol Cell Neurosci. 2020;106:103502.
120. Schoof M, Hellwig M, Harrison L, Holdhof D, Lauffer MC, Niesen J, et al. The basic helix-loop-helix transcription factor TCF4 impacts brain architecture as well as neuronal morphology and differentiation. Eur J Neurosci. 2020;51(11):2219-35.
121. Hill MJ, Killick R, Navarrete K, Maruszak A, McLaughlin GM, Williams BP, et al. Knockdown of the schizophrenia susceptibility gene TCF4 alters gene expression and proliferation of progenitor cells from the developing human neocortex. J Psychiatry Neurosci. 2017;42(3):181-8.
122. Kennedy AJ, Rahn EJ, Paulukaitis BS, Savell KE, Kordasiewicz HB, Wang J, et al. Tcf4 Regulates Synaptic Plasticity, DNA Methylation, and Memory Function. Cell Rep. 2016;16(10):2666-85.
123. Adachi M, Lin PY, Pranav H, Monteggia LM. Postnatal Loss of Mef2c Results in Dissociation of Effects on Synapse Number and Learning and Memory. Biol Psychiatry. 2016;80(2):140-8.
124. Mitchell AC, Javidfar B, Pothula V, Ibi D, Shen EY, Peter CJ, et al. MEF2C transcription factor is associated with the genetic and epigenetic risk architecture of schizophrenia and improves cognition in mice. Mol Psychiatry. 2018;23(1):123-32.
125. Desfilis E, Abellan A, Sentandreu V, Medina L. Expression of regulatory genes in the embryonic brain of a lizard and implications for understanding pallial organization and evolution. J Comp Neurol. 2018;526(1):166-202.
126. Bruce LL, Neary TJ. The limbic system of tetrapods: a comparative analysis of cortical and amygdalar populations. Brain Behav Evol. 1995;46(4-5):224-34.
127. Chen M, Maimaitili M, Habekost M, Gill KP, Mermet-Joret N, Nabavi S, et al. Rapid generation of regionally specified CNS neurons by sequential patterning and conversion of human induced pluripotent stem cells. Stem Cell Res. 2020;48:101945.
128. Zhang Y, Pak C, Han Y, Ahlenius H, Zhang Z, Chanda S, et al. Rapid single-step induction of functional neurons from human pluripotent stem cells. Neuron. 2013;78(5):785-98.
129. Hebert JM, Mishina Y, McConnell SK. BMP signaling is required locally to pattern the dorsal telencephalic midline. Neuron. 2002;35(6):1029-41.

130. Medina L, Abellan A, Desfilis E. Contribution of Genoarchitecture to Understanding Hippocampal Evolution and Development. Brain Behav Evol. 2017;90(1):25-40.
131. Calvo SE, Pagliarini DJ, Mootha VK. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. Proc Natl Acad Sci U S A. 2009;106(18):7507-12.
132. Dinman JD. Mechanisms and implications of programmed translational frameshifting. Wiley Interdiscip Rev RNA. 2012;3(5):661-73.
133. Wang X, Hou J, Quedenau C, Chen W. Pervasive isoform-specific translational regulation via alternative transcription start sites in mammals. Mol Syst Biol. 2016;12(7):875.
134. Pesold C, Impagnatiello F, Pisu MG, Uzunov DP, Costa E, Guidotti A, et al. Reelin is preferentially expressed in neurons synthesizing gamma-aminobutyric acid in cortex and hippocampus of adult rats. Proc Natl Acad Sci U S A. 1998;95(6):3221-6.
135. Ramsden HL, Surmeli G, McDonagh SG, Nolan MF. Laminar and dorsoventral molecular organization of the medial entorhinal cortex revealed by large-scale anatomical analysis of gene expression. PLoS Comput Biol. 2015;11(1):e1004032.
136. Polepalli JS, Gooch H, Sah P. Diversity of interneurons in the lateral and basal amygdala. NPJ Sci Learn. 2020;5:10.
137. Miao C, Cao Q, Moser MB, Moser EI. Parvalbumin and Somatostatin Interneurons Control Different Space-Coding Networks in the Medial Entorhinal Cortex. Cell. 2017;171(3):507-21 e17.
138. Krishnaswami SR, Grindberg RV, Novotny M, Venepally P, Lacar B, Bhutani K, et al. Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. Nat Protoc. 2016;11(3):499-524.
139. Pfisterer U, Petukhov V, Demharter S, Meichsner J, Thompson JJ, Batiuk MY, et al. Identification of epilepsy-associated neuronal subtypes and gene expression underlying epileptogenesis. Nat Commun. 2020;11(1):5038.
140. Gill KP, Denham M. Optimized Transgene Delivery Using Third-Generation Lentiviruses. Curr Protoc Mol Biol. 2020;133(1):e125.
141. Thomsen R, Solvsten CA, Linnet TE, Blechingberg J, Nielsen AL. Analysis of qPCR data by converting exponentially related Ct values into linearly related X0 values. J Bioinform Comput Biol. 2010;8(5):885-900.
142. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.
143. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. Nat Biotechnol. 2015;33(5):495-502.
144. van der Maaten L, Hinton G. Visualizing data using t-SNE. Journal of Machine Learning Research. 2008;9:2579-605.
145. Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. J Stat Mech. 2008:P10008.
146. Fan X, Dong J, Zhong S, Wei Y, Wu Q, Yan L, et al. Spatial transcriptomic survey of human embryonic cerebral cortex by single-cell RNA-seq analysis. Cell Res. 2018;28(7):730-45.
147. Astick M, Vanderhaeghen P. From Human Pluripotent Stem Cells to Cortical Circuits. Curr Top Dev Biol. 2018;129:67-98.
148. 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.

149. Tirosh I, Izar B, Prakadan SM, Wadsworth MH, 2nd, Treacy D, Trombetta JJ, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science. 2016;352(6282):189-96.

Figure legends

Figure 1. Single-cell profiling of the entorhinal cortex reveals thirty-two distinct cell clusters of major cell type populations. (A) The computational pipeline includes ten batches of single-cell and single-nuclei isolated cells from E50, E60, E70 and P75 which were captured using the 10X genomics platform. Batches were normalized using the Fast MNN approach and dimensional reduction and clustering were performed in Seurat followed by subclustering and analyses of selected clusters. **(B)** A t-SNE plot of 24,294 cells merged from all timepoints revealed thirty-two distinct populations following FastMNN batch correction. Oligodendro-glia/-cytes (Oligo), astrocytes and radial glia (Astro & RGC), intermediate progenitors (IP), excitatory neurons (Excitatory), interneurons (IN) and microglia (Microglia) **(C)** tSNE plot annotated by batch demonstrates good integration of subtypes within same ages following correction by FastMNN of the 10 datasets (pre-batch correction shown in Figure S5C). **(D)** Analysis of canonical marker genes resulted in categorization of several Oligo, Astro & RGC, IP, Excitatory, IN and Microglia clusters. The number (n) of cells, mean UMI counts (unique transcripts) and mean number of different expressed genes for each cluster are also represented. Error bars denote the SD. **(E)** Projection of the dataset onto an already-annotated human fetal prefrontal cortex and human medial temporal gyrus single-cell dataset consolidates marker gene-driven annotation of the dataset. Abbreviations: NPC, neural progenitor cells; HBB, marker expressed in human blood cells.

Figure 2. Subclustering of entorhinal cortex (EC) interneurons (INs) reveals ten clusters with unique gene signatures. (A) A t-SNE plot of the subclustering identifies ten interneuron (IN) clusters with one cluster more transcriptionally distinct than the others. IN clusters highlighted in yellow in the inset image were subclustered from the parent dataset and are depicted in the top left corner. **(B)** An analysis of the developmental stages across the ten clusters highlights IN1, IN5, IN7, IN8 and IN9 are found exclusively during development and suggestive of IN progenitors and the remaining clusters are adult INs. **(C)** Unique gene signatures were identified for the ten clusters by differential gene expression analysis, together with the population specific expression of canonical markers for interneurons (IN) derived both from the caudal ganglionic eminence (CGE) and the medial ganglionic eminence (MGE). Localization of genes with names in blue are shown in (D). * marks the human ortholog names (Table S1). **(D)** Sagittal sections from the ISH datasets from the Allen Developing Mouse Brain Atlas and Allen Mouse Brain Atlas highlight the location of selected markers from the IN clusters across the EC layers in the developing and postnatal mouse brain (84-86). Image credit: Allen Institute. Scale bars: 250 μm (for E18.5, P4 and P14) and 500 μm (for P56).

Figure 3. Subclustering of oligodendro-cytes and -glia reveals oligodendrocyte progenitor cells (OPCs) emerge in the early entorhinal cortex. (A) Visualization of the oligodendroglia subset (number (n) of cells = 2,696) shows five clusters in a t-SNE plot. Oligodendroglia clusters highlighted in orange in the inset image were subclustered from the parent dataset and are depicted in the top left corner. **(B)** The developmental stages of the oligodendroglia visualized by a bar plot highlight a distinct adult and fetal populations and populations that exist both during development as early as E50 and postnatally. **(C)** Canonical lineage markers and

unique gene signatures for the five clusters were identified by differential gene expression analysis. Localization of genes with names in blue is shown in (E). * marks the human ortholog names (Table S1). **(D)** Distribution of cells within different cell cycle phases highlights OLI4 is largely a non-mitotic population and the remaining clusters are mitotically active. Note, G1-phase is indistinguishable from the G0-phase. **(E)** Sagittal sections from the ISH datasets from the Allen Developing Mouse Brain Atlas and Allen Mouse Brain Atlas (84-86) highlight the location of selected markers from the OPC and OLI subtypes. Image credit: Allen Institute. Scale bars: 250 μm (for E18.5 and P4) and 500 μm (for P28 and P56).

Figure 4. Subclustering of the excitatory neurons and intermediate progenitors of the developing entorhinal cortex (EC) highlight three *RELN*+ clusters and five pyramidal neuron clusters. (A) A t-SNE plot highlights subclustering of excitatory neurons and intermediate progenitors (number (n) of cells = 6,833) results in ten distinct clusters. The inserted image shows the excitatory neuron clusters highlighted in blue and intermediate progenitors in green, that were subclustered from the parent dataset. **(B)** Distribution of developmental stages within the ten clusters highlights five populations persist postnatally (three pyramidal neurons, PYR0, PYR2, PYR5 and two Reelin+ cell populations, RELN6, RELN7). Three IP populations were identified IP4, IP8 and RELN9 which the latter also expressed IN markers and RELN. IP4 was also detected postnatally but expressed IP makers *EOMES, PAX6* and *SOX2* **(C)** Unique gene signatures for the ten clusters were determined following differential gene expression analysis. Localization of genes with names in blue are shown in (D). * marks the human ortholog names (Table S1). **(D)** Sagittal sections from the ISH datasets from the Allen Developing Mouse Brain Atlas and Allen Mouse Brain Atlas highlight the location of selected markers from the RELN and PYR clusters across the EC layers at different developmental stages and in the adult (84-86). Image credit: Allen Institute. Scale bars: 250 μm (for E18.5 and P4) and 500 μm (for P56).

Figure 5. Unique transcriptional factors upregulated in *RELN* positive clusters. (A) Expression of the top twenty enriched transcription factors (TFs, x-axis) were identified for each of the *RELN*-positive and putative stellate cell (SC) populations (RELN6, 7) and are visualized across the identified progenitor and excitatory neuron populations (y-axis) in the developing EC. The selected genes for reprogramming induced pluripotent stem cells (iPSCs) are highlighted in blue within the RELN7 cluster. **(B)** The laminar expression of the six SC highlighted TFs (*Runx1t1, Eya2, Foxp1, Sox5, Tcf4*, and *Mef2c*) was investigated by ISH in sagittal sections of the publically available Allen Mouse Brain Atlas (84, 85). Image credit: Allen Institute. The red dotted lines demarcate the *lamina dissecans*. **(C)** Schematic diagram of the protocol used for directly reprogramming iPSCs into SCs using the selected TFs, including timing of media and matrix change, transfection, gene induction (using DOX), addition of patterning factors (CHIR 99021 and BMP4), and collection of samples. Abbreviations: Doxycycline (DOX); pluripotent stem cell media (mTESR); neurobasal medium with medial pallium patterning factors (MP patterning); neurobasal medium for maintenance of neurons (NBM).

Figure 6. Direct reprogramming of induced pluripotent stem cells (iPSCs) results in the production of stellate cell-like cells. (A) The morphological changes of the two iPSC lines SFC 180-01-01 (SFC) and WTSli024A (WTS) are shown after transduction with the stellate cell transcription factors alone (SC TFs), the SC TFs plus *Ngn2*, *Ngn2* alone and without any TFs at Day (D)10 and D20 post transduction. Scale bar: 100 μm. (B) RT-qPCR analysis of pooled D10 and D20 samples highlights upregulated expression of *RELN, SATB2, CALB1*

and *LEF1* in SC TF transduced iPSCs compared with iPSCs not transduced with any TFs. Expression values were normalized to *GAPDH* and presented as fold-change relative to No TFs. Bars represent the mean values of three replicates. The error bars represent the SD. Statistical one-way ANOVA, with Tukey's multiple comparisons test was performed and significance is reported as: *p < 0.05, **p < 0.01, ***p <0.001. **(C)** Expression of BCL11B was determined in the transduced SFC and WTS iPSCs, including the SC TFs alone treatment, at D20. HO; Hoechst. Scale bar: 100 μm.

Figure S1. Mitochondrial gene expression in the whole scRNA-seq dataset, evaluation of interneuron (IN) genes in the entorhinal cortex (EC) excitatory neuron/intermediate progenitor (IP) clusters. (A) The top ten gene features of cluster 31 identified in Seurat using FindAllMarkers (with Wilcoxon Rank Sum test) showed enrichment of mitochondrial genes in the cluster, likely representing dying cells, which resulted in the cluster being removed from further analyses. **(B)** Expression of the IN markers, *GAD1* and *GAD2* in the IP and excitatory neuron clusters highlight low expression across all the subtypes, with the exception of *GAD2* in PYR12 which suggests it may be an IN progenitor.

Figure S2. Expression of RELN, PYR and IP cluster markers in saggital sections of the entorhinal cortex (EC) from the Allen Developing Mouse Brain Atlas and Allen Mouse Brain Atlas. Expression of select markers for varying excitatory neuron clusters highlighted in Figure 4C were investigated in ISH sagittal sections revealing putative locations of the cell types, such as RELN7 and PYR2 neurons residing in the superficial layers. Image credit: Allen Institute (82-84). Scale bars: 250 μm (for E18.5 and P4) and 500 μm (for P14 and P56).

Figure S3. Direct programming of induced pluripotent stem cells (iPSCs) using a leave-one-out approach of the stellate cell-specific transcription factors in the presence of *Ngn2* resulted in the production of stellate cell-like cells. (A) Morphological changes of the two iPSC lines, SFC 180-01-01 (SFC) and WTSli024A (WTS) following transduction with different combinations of five out of the six stellate cell transcription factors (SC TFs), together with *Ngn2* at day (D)10 and D20. Scale bar: 100 μm. **(B)** Expression of BCL11B in the SFC and WTS cell lines at D10 after transduction with the SC TFs alone, the SC TFs with Ngn2, Ngn2 alone, or with no TFs. HO; Hoechst. Scale bar: 100 μm.

Figure S4. BCL11B expression in induced pluripotent stem cells (iPSCs) following different combinations of induced transcription factor expression. BCL11B expression was observed in both iPSC lines SFC 180-01-01 (SFC) and WTSli024A (WTS) transduced with different combinations of five stellate cell transcription factors (SC TFs) in the presence of *Ngn2*, at Day (D)10 and 20 post transduction. HO; Hoechst. Scale bar: 100 μm.

Figure S5. Validation data for the scRNA-seq experiments. (A) Violin plots of the number (n) of UMIs and genes in each single-cell sequencing batch. **(B)** Scatter plots of the number (n) of UMIs versus genes for each single-cell sequencing batch. **(C)** Fraction of transcripts out of total transcripts for a given gene in empty droplets at representative stages (E70 and Adult), for both isolated nuclei (nc) and whole cell samples. Genes with a fraction above 0.30 were not included as variable genes in the downstream analysis. Abbreviations: n, number; E, embryonic **(D)** A t-SNE plot of the scRNA-seq dataset prior to fastMNN batch correction.

Supplementary Information

Additional file 1: Supplementary Figures
Figures S1-S5 with figure legends

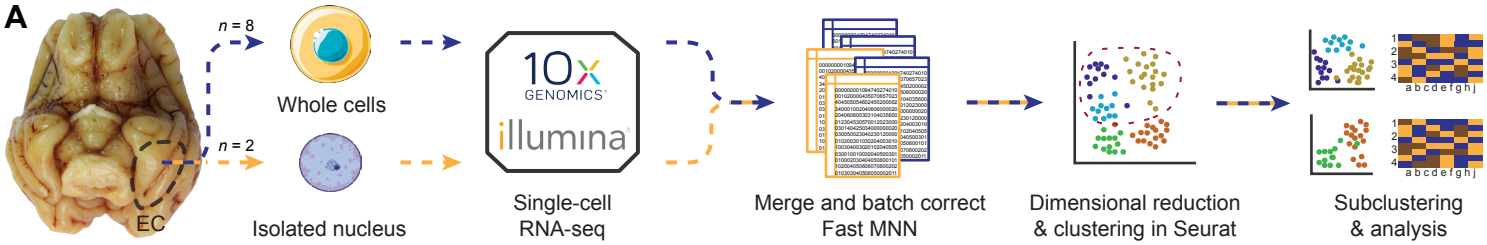
Additional file 2: Table S1
Table S1: Human ortholog gene names identified from pig ensemble names

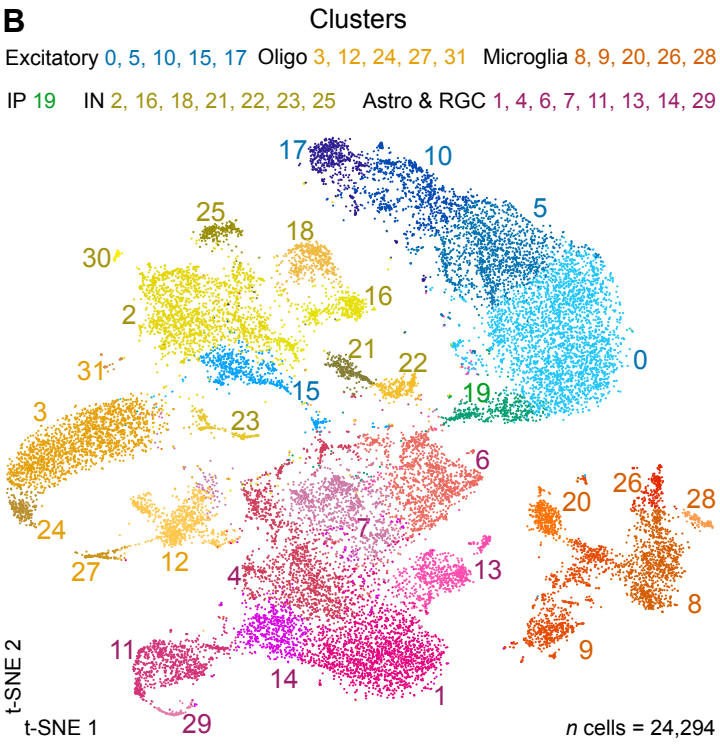
Additional file 3: Table S2
Table S2: Primers for cloning of plasmids

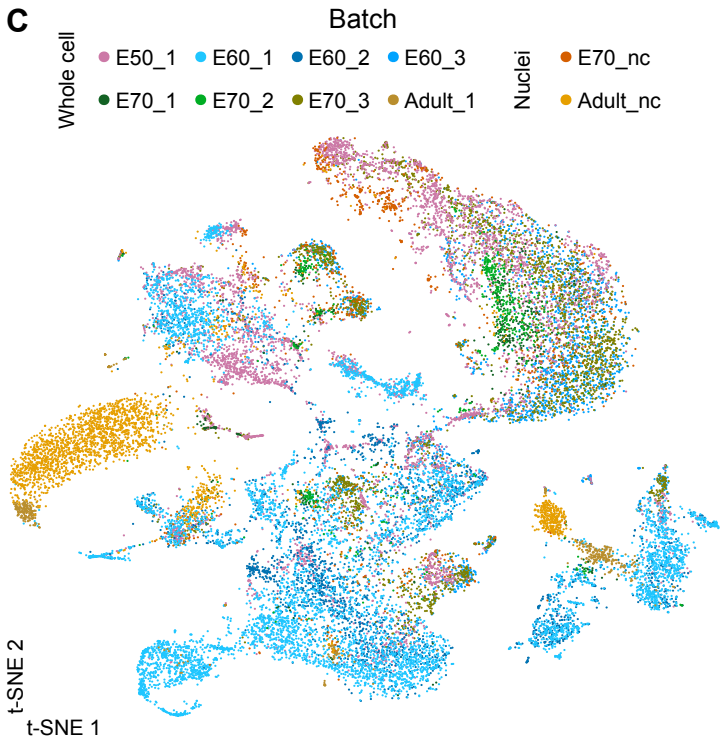
Additional file 4: Table S3
Table S3: Primers for qPCR

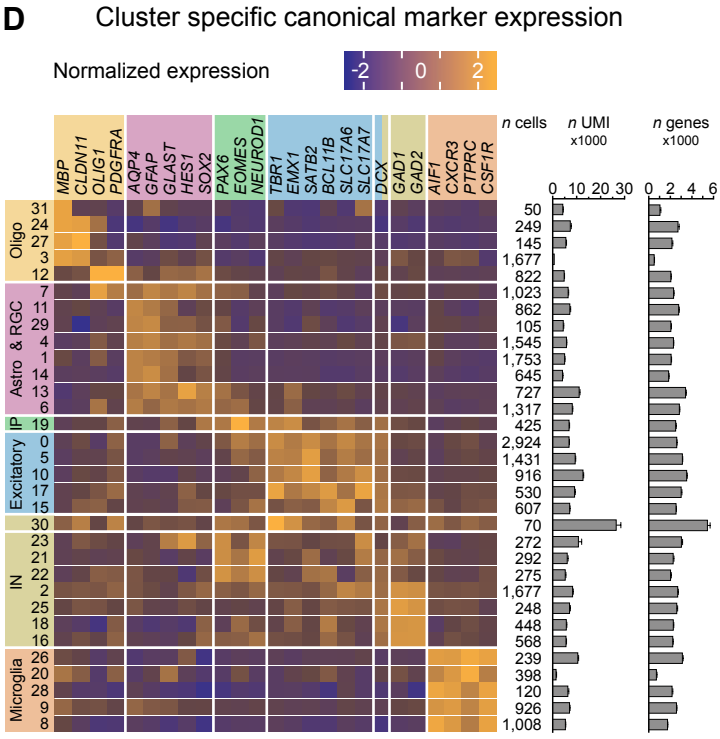
Additional file 5: Table S4

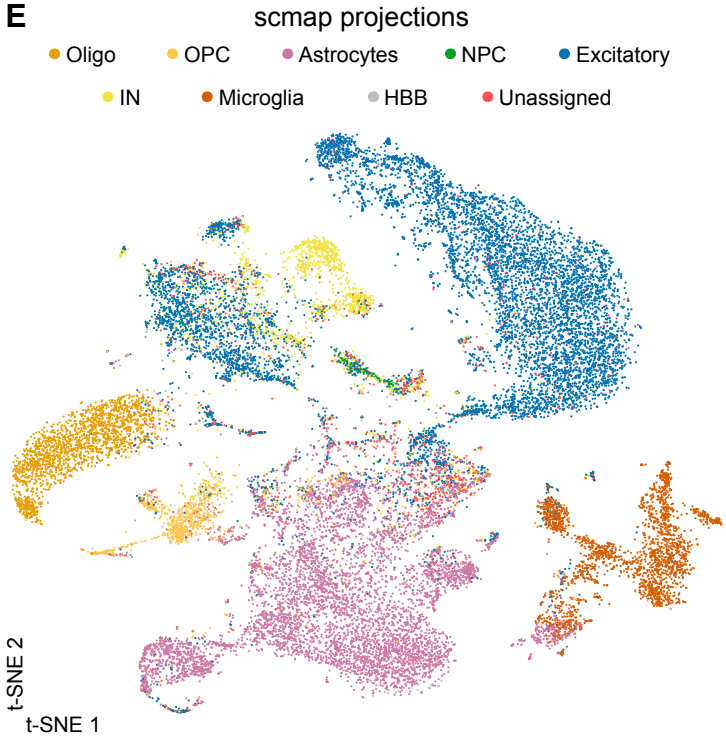
Table S4: Statistics for scRNA-seq libraries

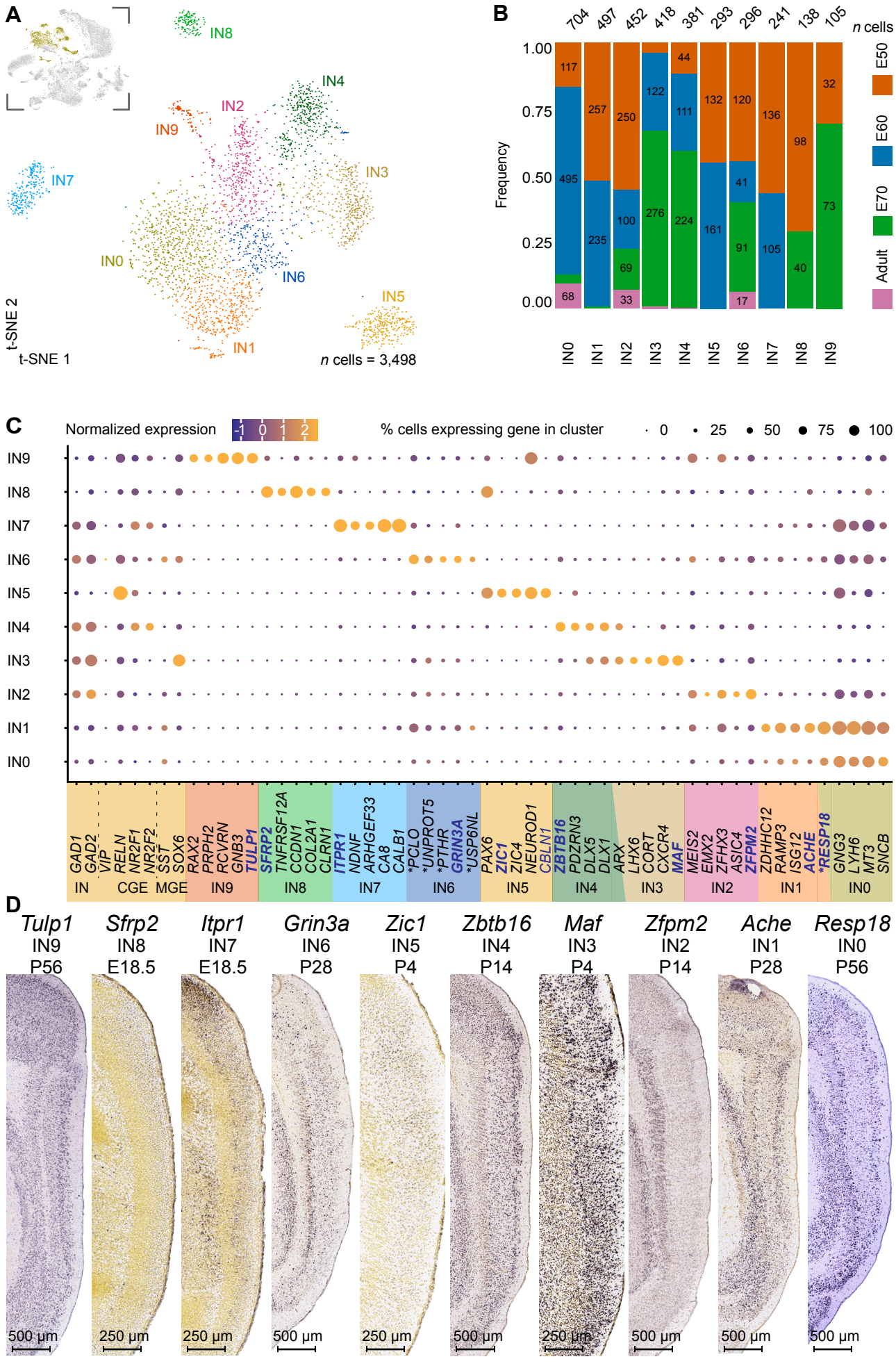


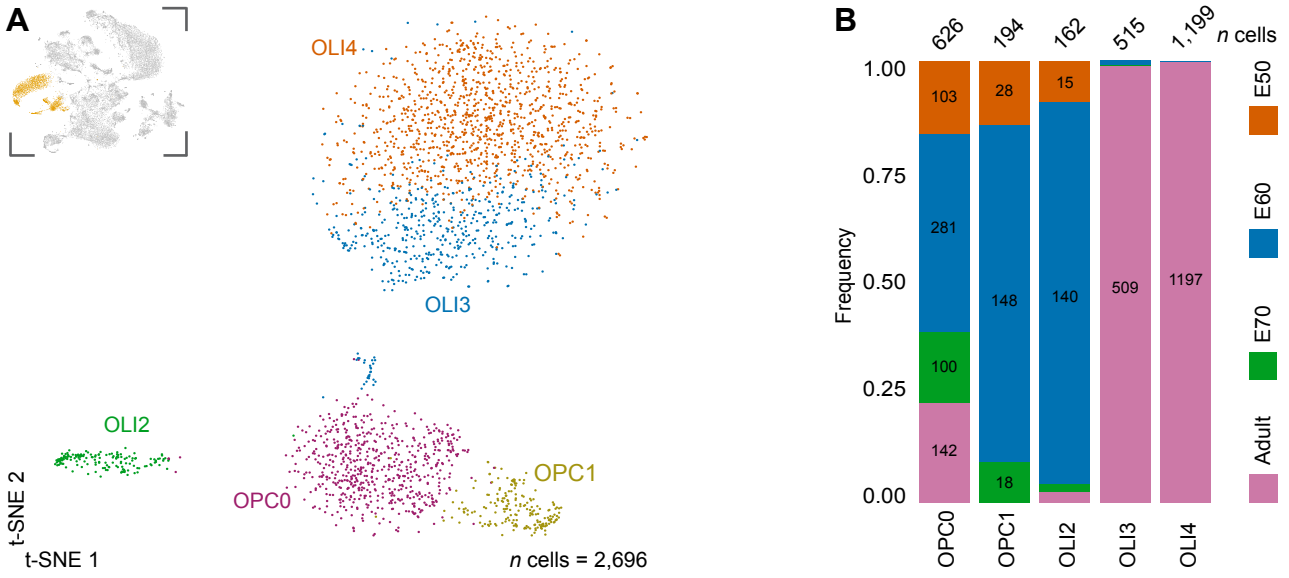


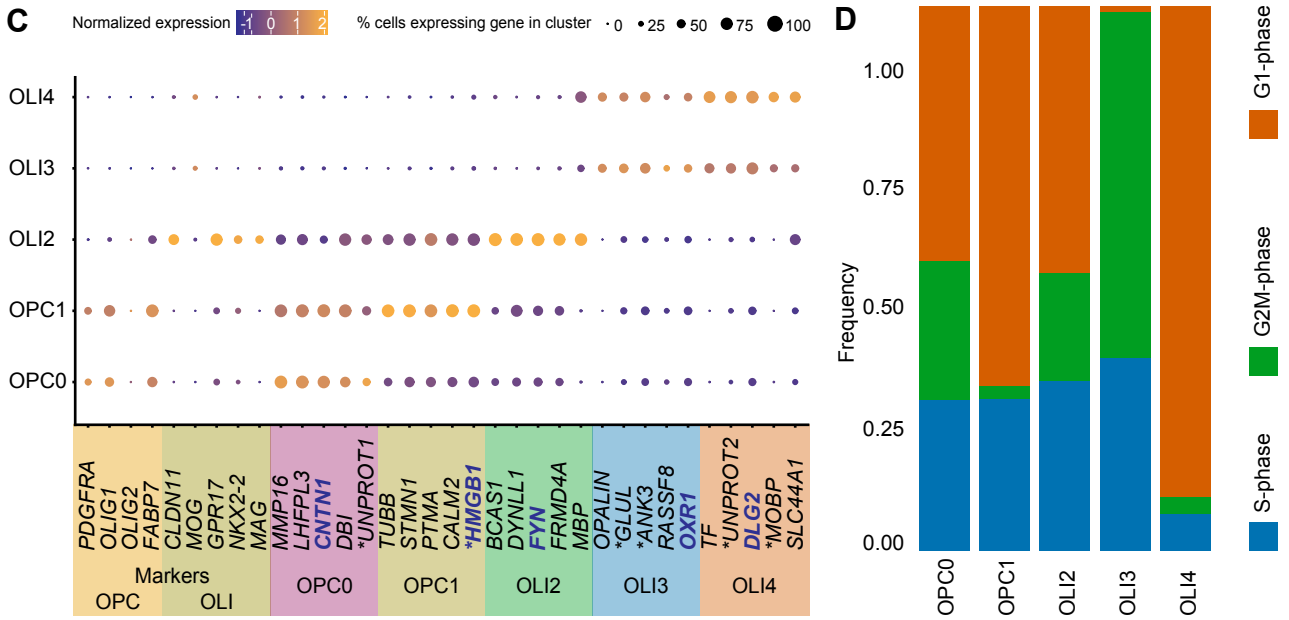


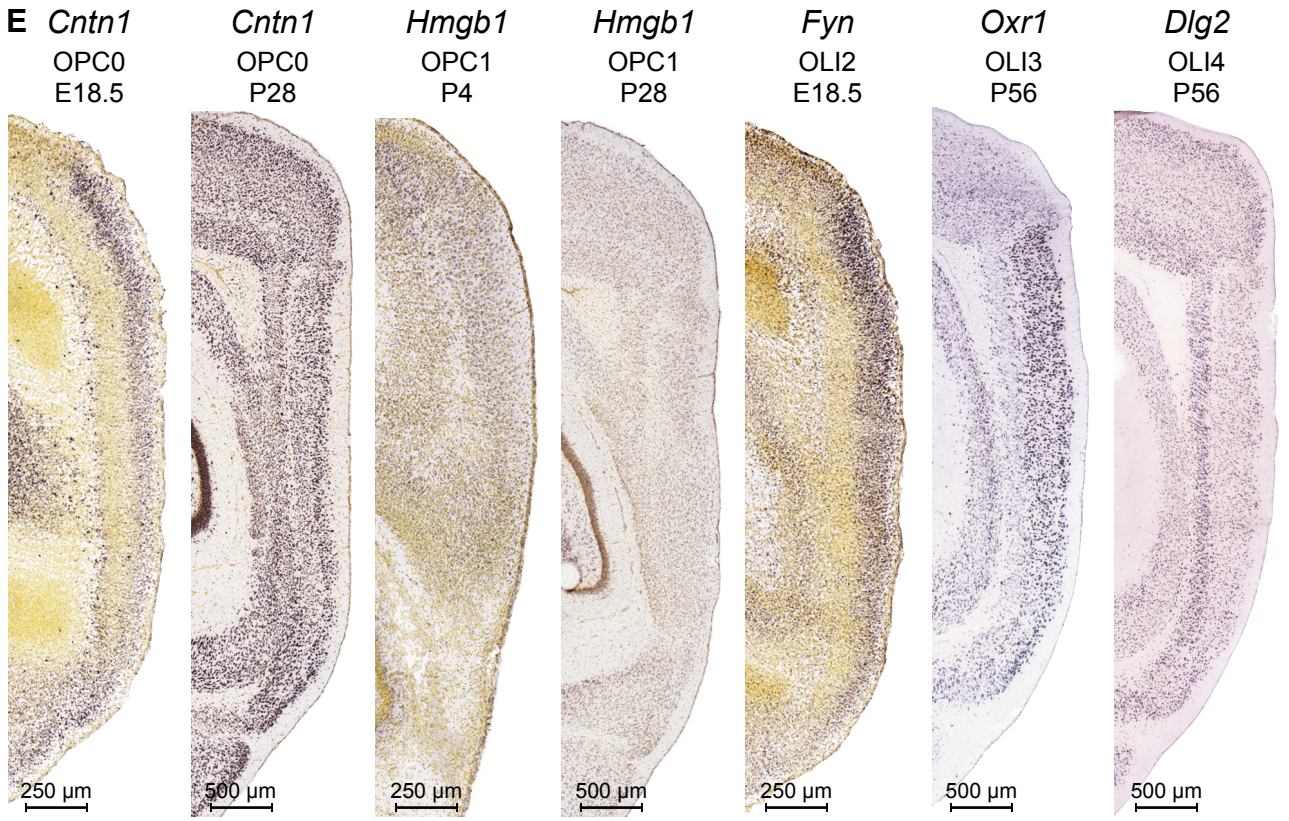


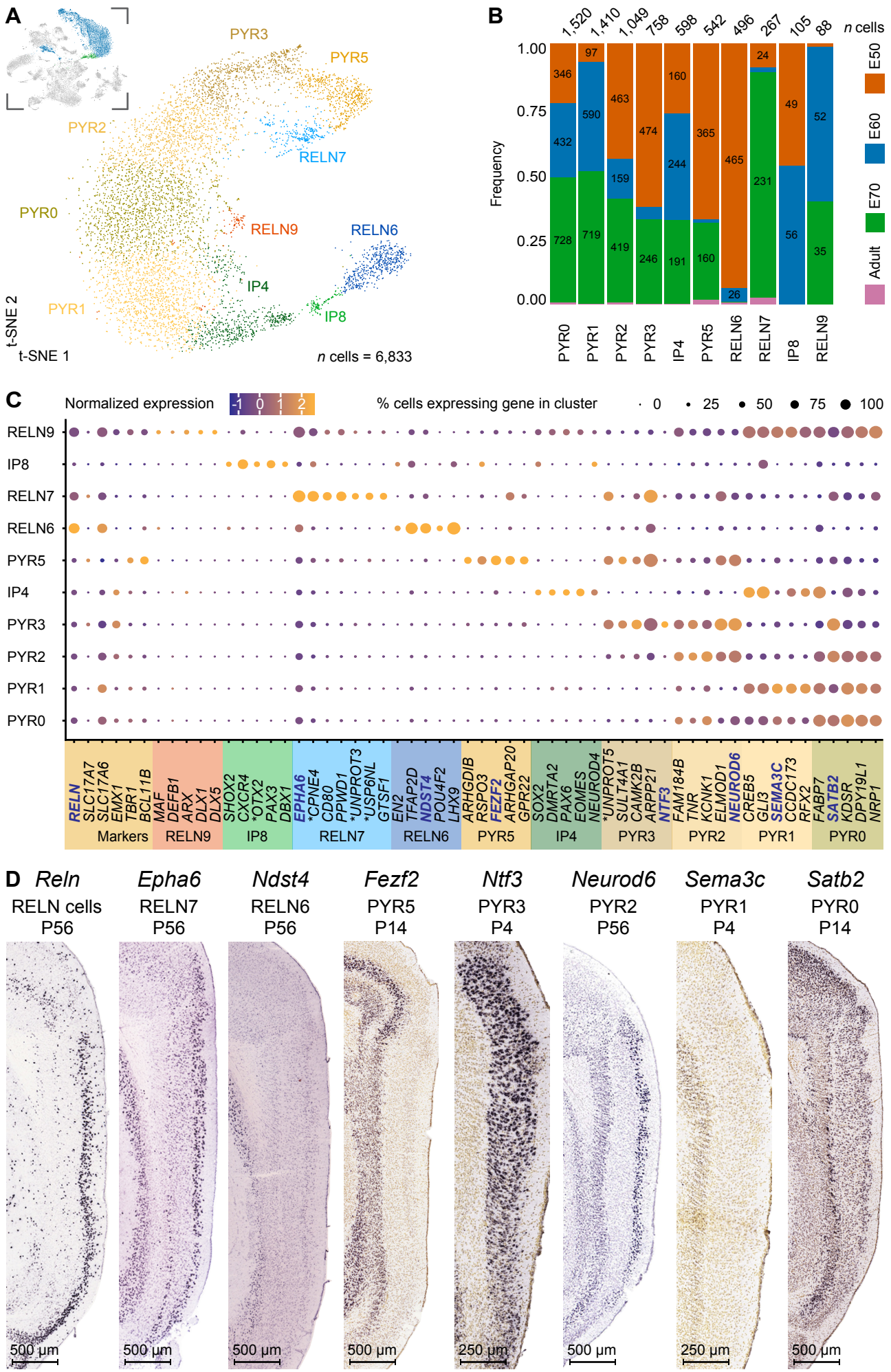



A
IN8
IN4
IN2
IN9
IN3
IN7
IN0
IN6
IN5
IN1
t-SNE 2
t-SNE 1
n cells = 3,498
B
n cells
704 497 452 418 381 293 296 241 138 105
Frequency
1.00
0.75
0.50
0.25
0.00
117 495 68
257 235
250 100 69 33
122 276
44 111 224
132 161
120 41 91 17
136 105
98 40
32 73
E50
E60
E70
Adult
IN0 IN1 IN2 IN3 IN4 IN5 IN6 IN7 IN8 IN9
C
Normalized expression
-1 0 1 2
% cells expressing gene in cluster
0 25 50 75 100
IN9
IN8
IN7
IN6
IN5
IN4
IN3
IN2
IN1
IN0
GAD1 GAD2 VIP RELN NR2F1 NR2F2 SST SOX6 RAX2 PRPH2 RCVRN GNB3 TULP1
SFRP2 TNFRSF12A CCDN1 COL2A1 CLRN1
ITPR1 NDNF ARHGEF33 CA8 CALB1
*PCLO *UNPROT5 *PTHR GRIN3A *USP6NL
PAX6 ZIC1 ZIC4 NEUROD1 CBLN1
ZBTB16 PDZRN3 DLX5 DLX1 ARX
LHX6 CORT CXCR4 MAF
MEIS2 EMX2 ZFHX3 ASIC4 ZFPM2
ZDHHC12 RAMP3 ISG12 ACHE *RESP18
GNG3 LYH6 MT3 SNCB
IN CGE MGE IN9 IN8 IN7 IN6 IN5 IN4 IN3 IN2 IN1 IN0
D
Tulp1 IN9 P56
Sfrp2 IN8 E18.5
Itpr1 IN7 E18.5
Grin3a IN6 P28
Zic1 IN5 P4
Zbtb16 IN4 P14
Maf IN3 P4
Zfpm2 IN2 P14
Ache IN1 P28
Resp18 IN0 P56
500 µm
250 µm
250 µm
500 µm
250 µm
500 µm
250 µm
500 µm
500 µm
500 µm








A
PYR3
PYR5
PYR2
RELN7
PYR0
RELN9
RELN6
IP4
PYR1
IP8
t-SNE 2
t-SNE 1
n cells = 6,833
B
n cells
1,520
1,410
1,049
758
598
542
496
267
105
88
Frequency
1.00
0.75
0.50
0.25
0.00
346
432
728
97
590
719
463
159
419
474
246
160
244
191
365
160
465
26
24
231
49
56
52
35
E50
E60
E70
Adult
PYR0
PYR1
PYR2
PYR3
IP4
PYR5
RELN6
RELN7
IP8
RELN9
C
Normalized expression
-1
0
1
2
% cells expressing gene in cluster
0
25
50
75
100
RELN9
IP8
RELN7
RELN6
PYR5
IP4
PYR3
PYR2
PYR1
PYR0
RELN SLC17A7 SLC17A6 EMX1 TBR1 BCL11B
MAF DEFB1 ARX DLX1 DLX5
SHOX2 CXCR4 *OTX2 PAX3 DBX1
EPHA6 *CPNE4 CD80 PPWD1 *UNPROT3 *USP6NL GTSF1
EN2 TFAP2D NDST4 POU4F2 LHX9
ARHGDIB RSPO3 FEZF2 ARHGAP20 GPR22
SOX2 DMRTA2 PAX6 EOMES NEUROD4
*UNPROT5 SULT4A1 CAMK2B ARPP21 NTF3
FAM184B TNR KCNK1 ELMOD1 NEUROD6
CREB5 GLI3 SEMA3C CCDC173 RFX2
FABP7 SATB2 KDSR DPY19L1 NRP1
Markers
RELN9
IP8
RELN7
RELN6
PYR5
IP4
PYR3
PYR2
PYR1
PYR0
D
Reln
RELN cells
P56
Epha6
RELN7
P56
Ndst4
RELN6
P56
Fezf2
PYR5
P14
Ntf3
PYR3
P4
Neurod6
PYR2
P56
Sema3c
PYR1
P4
Satb2
PYR0
P14
500 µm
500 µm
500 µm
500 µm
250 µm
500 µm
250 µm
500 µm