Rapid and robust directed differentiation of mouse epiblast stem cells into definitive endoderm and forebrain organoids

Daniel Medina-Cano[1,2,*], Emily K. Corrigan[1,2,3,*], Rachel Glenn[1,2,4,&], Mohammed Islam[1,2,&], Yuan Lin[1,2], Juliet Kim[1,2], Hyunwoo Cho[1,2], Thomas Vierbuchen[1,2,#]

Affiliations:
[1] Developmental Biology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, USA
[2] Center for Stem Cell Biology, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, USA
[3] Present address: Developmental and Stem Cell Biology Graduate Program, University of California San Francisco, San Francisco, CA, USA
[4] Cell and Developmental Biology Program, Weill Cornell Graduate School of Medical Sciences, Cornell University, New York, NY, USA
[*]These authors contributed equally
[&] These authors contributed equally
[#] For correspondence: vierbuct@mskcc.org

SUMMARY STATEMENT

New optimized protocols for directed differentiation of mouse epiblast stem cells into definitive endoderm and forebrain-patterned organoids.

ABSTRACT

Directed differentiation of pluripotent stem cells (PSCs) is a powerful model system for deconstructing embryonic development. Although mice are the most advanced mammalian model system for genetic studies of embryonic development, state-of-the-art protocols for directed differentiation of mouse PSCs into defined lineages tend to be slower and generate target cell types with lower purity than analogous protocols for human PSCs, limiting their application as models for mechanistic studies of development. Here, we examine the potential of mouse epiblast stem cells (EpiSCs) cultured in media containing Wnt pathway inhibitors (primed ground state conditions) as a starting point for directed differentiation. As a proof-of-concept, we focused our efforts on two specific cell/tissue types that have proven difficult to generate efficiently and reproducibly from mouse embryonic stem cells: definitive endoderm and neural organoids. First, we developed a new protocol that can rapidly generate nearly pure definitive endoderm from EpiSCs. Second, we developed a protocol for generating forebrain organoids that model the development of prethalamic and hippocampal neurons. These significantly improved differentiation models present new possibilities for combining mouse genetic tools and resources with *in vitro* differentiation to characterize the molecular and cellular mechanisms of embryonic development.

KEY WORDS: pluripotent stem cells, directed differentiation, definitive endoderm, forebrain development, neural organoids

INTRODUCTION

Over the last two decades, methods for the directed differentiation of human PSCs (hPSCs) into numerous lineages have been developed and refined. As a result, it is now possible to generate nearly pure populations of multiple specific cell types or 3D multi-cellular organoids that resemble developing human tissues (Huch et al., 2017; Loh et al., 2016; Tchieu et al., 2017; Yiangou et al., 2018). However, while hPSC-based models for studying development advanced, analogous protocols for mouse PSCs have not yet emerged. Current state-of-the-art protocols for directed differentiation of mouse PSCs tend to be slower and more complex than equivalent hPSC protocols, and they often fail to generate target cell types with high purity (Ortmann et al., 2020). Robust and efficient protocols for directed differentiation of mouse PSCs into defined lineages would be synergistic with *in vivo* studies in mouse and with studies of hPSC-based models of human embryonic development.

A critical difference between most mouse and human directed differentiation protocols is the starting cell state. Human PSCs are generally cultured in the primed pluripotent state, which is roughly equivalent to cells of the pluripotent epiblast around the start of gastrulation (~E6.5 in mouse) (Loh et al., 2015; Rossant and Tam, 2017). In contrast, mouse PSCs are predominantly grown in the naïve pluripotent state (mouse embryonic stem cells), which is equivalent to cells in the inner cell mass of the pre-implantation embryo at ~E3.5-4.0 (Nichols and Smith, 2012). Like their embryonic equivalents, naïve PSCs must first exit the naïve pluripotent state before they can properly respond to differentiation cues (Morgani et al., 2017; Smith, 2017). As a result, mouse naïve PSC directed differentiation protocols must include an additional first step of ~48 hours in which cells exit the naïve pluripotent state and transition into an early formative and/or later primed pluripotent state that is competent to respond to signals that induce differentiation into specific somatic lineages (Smith, 2017). The requirement for this additional initial step is problematic because the naïve-to-primed transition is inherently asynchronous, leading to a heterogeneous starting population of formative/primed cells before differentiation into a specific germ layer has even commenced (Kalkan et al., 2017; Strawbridge et al., 2020). The naïve-primed conversion step thus represents an impediment to the development of rapid and efficient protocols for directed differentiation that use mouse naïve PSCs as a starting population.

Mouse PSCs can also be cultured in a primed pluripotent state, referred to as epiblast stem cells (EpiSCs). EpiSCs share many similarities with human PSCs, including the ability to respond rapidly to differentiation stimuli (Brons et al., 2007; Tesar et al., 2007; Vallier et al., 2009; Weinberger et al., 2016). Protocols that use EpiSCs as a starting point for directed differentiation have been developed, including robust protocols for generation of oligodendrocyte progenitor cells and neuromesodermal progenitors (Edri et al., 2019; Najm et al., 2011). However, EpiSC-based protocols have not been widely adopted, likely for several reasons, including, 1) EpiSCs are prone to spontaneous differentiation and thus are more difficult to culture than mouse naïve PSCs or human PSCs, 2) EpiSCs have been reported to exhibit significant line-to-line variability in their differentiation propensity, and 3) EpiSCs express markers of lineage-primed epiblast cells transiting through the anterior primitive streak (T, FOXA2), which may limit their functional potential to differentiate into lineages derived from other regions of the primitive streak (Bernemann et al., 2011; Jouneau, 2019; Kojima et al., 2014; Kurek et al., 2015; Rossant, 2015; Rossant and Tam, 2017; Song et al., 2016; Tsakiridis et al., 2014; Weinberger et al., 2016).

More recently, several groups demonstrated that inhibition of the canonical Wnt signaling pathway can limit spontaneous differentiation in EpiSC cultures, enabling long-term maintenance of more homogeneous EpiSC cultures (Kurek et al., 2015; Sumi et al., 2013; Wu et al., 2015). Despite this progress, EpiSCs cultured in these conditions (primed ground state) are still rarely used for directed differentiation experiments (Edri et al., 2019; Morgani et al., 2017; Najm et al., 2011). Therefore, whether primed ground state EpiSCs have the potential to improve directed differentiation towards multiple lineages in vitro remains a critical gap in the field.

Here, we examine the potential of ground-state mouse EpiSCs as a starting point for *in vitro* directed differentiation studies. We focused on two specific cell populations that have proven difficult to generate reproducibly and efficiently from mouse naïve PSCs: definitive endoderm (DE) and neural organoids. First, we systematically optimized conditions to differentiate EpiSCs into DE, developing a protocol that converts EpiSCs into >90% pure DE in only 40 hours. Second, we show that EpiSCs can robustly and reproducibly generate forebrain-patterned neural organoids. In depth

characterization of forebrain-patterned organoids revealed that they undergo polarization into multiple distinct domains of progenitors, including pre-thalamic (P3) region of the caudal forebrain (diencephalon), hippocampus and cortical hem, as well as oligodendrocyte progenitor cells. The prethalamic progenitors, which gives rise primarily to inhibitory (GABAergic) neurons of the thalamic reticular nucleus and the zona incerta, are notable because they have not been generated efficiently in previous mouse and human thalamic organoids (Shiraishi et al., 2017; Xiang et al., 2019). Taken together, our data suggest that ground-state EpiSCs can be a powerful platform for *in vitro* directed differentiation studies in mouse, enabling comparative studies with mouse embryonic development *in vivo* and with hPSC directed differentiation models.

RESULTS

Derivation and characterization of a panel of ground state EpiSCs

We derived new primed ground state EpiSC lines from 12 distinct naïve PSC lines, including lines from several distinct inbred mouse strains: C57Bl/6J (n=5), B6129SF1/J (n=3), DBA/2J (n=1), and PWK/PhJ (n=3) (Czechanski et al., 2014) (**Supplementary Table 4**). Briefly, naïve PSCs were converted into epiblast like-cells (EpiLCs) using previously established protocols (Hayashi et al., 2011). Following EpiLC conversion, cells were split onto irradiated mouse embryonic fibroblast feeders and maintained EpiSC media containing a Tankyrase inhibitor (NVP-TNKS656) to block Wnt signaling (Kurek et al., 2015; Sumi et al., 2013; Wu et al., 2015). Consistent with previous studies, EpiSC lines cultured in these primed ground state conditions express markers of post-implantation epiblast (OCT4, SOX2, OTX2, NANOG), do not express markers specifically expressed in naïve pluripotent stem cells (KLF4), and exhibit sparse expression of markers of primed epiblast (FOXA2) and primitive streak (T) (**Fig. 1A-B**).

To characterize these EpiSC lines in greater depth, we measured gene expression (RNA-seq; n =12 lines) and mapped cis-regulatory elements across the genome (ATAC-seq; n = 9 lines). We observed high expression of primed (*Otx2*, *Oct6, Dnmt3b)* and general (*Oct4/Pou5f1, Sox2, Nanog*) pluripotency markers, and low expression of naïve pluripotency markers (*Klf4, Tbx3*, *Prdm14*, *Nr5a2*) across all lines, in agreement with results from immunofluorescence staining (**Fig. 1B-C**). Similarly, genes associated with lineage commitment/lineage priming (*T*, *Foxa2*, *Mixl1*, *Sox17, Gata6, Cdx2, Sox1)* exhibited low to moderate expression across all lines (**Fig. 1C**). Principle component analysis (PCA) of RNA-seq and ATAC-seq data revealed that genetic background is the primary driver of gene expression differences between lines (PC1, ~56% and ~89% of total variance, respectively, **Fig. 1D-E**). Individual EpiSC lines from the same genetic background generally had highly correlated profiles of both gene expression and chromatin accessibility, however, the C57Bl/6J EpiSC lines exhibited more variability in gene expression between lines that the other backgrounds (PC2, ~12% of total variance).

Next, we compared chromatin accessibility at cis-regulatory elements across the genome in our new EpiSC lines to cell populations in the developing embryo (epiblast, visceral endoderm, primitive streak, and germ layer progenitors) as well as EpiSC lines cultured in the presence (EpiSC- AFX) or absence (EpiSC-AF) of the Tankyrase inhibitor XAV939 (Kinoshita et al., 2020; Xiang et al., 2020). PCA of chromatin accessibility profiles revealed that our ground state EpiSC lines most closely resemble previously characterized EpiSC lines cultured in vitro (EpiSC-AF, EpiSC-AFX), followed by *in vivo* epiblast tissue, and they are least similar to primitive streak and lineage-committed progenitors (endoderm, mesoderm) (**Fig. 1F**). Taken together, these data suggest that EpiSCs cultured in primed ground state conditions exhibit expected features of primed PSCs, can be consistently derived via *in vitro* naïve-to-primed conversion of mouse naïve PSCs, and stably maintained in an undifferentiated state in culture.

Optimized directed differentiation of EpiSCs into definitive endoderm

We next sought to test whether EpiSCs cultured in primed ground state conditions can be directed to efficiently differentiate into DE. Current protocols for the directed differentiation of human PSCs into DE are rapid and efficient, and now serve as the basis for numerous robust protocols for generating progenitors of various endodermal organs (Yiangou et al., 2018). In contrast, state-of-the art protocols for the directed differentiation of mouse naïve PSCs into

DE take ~2 times longer than human protocols, are more complex, and generally fail to give rise to pure populations of DE (See meta-analyses in **Fig. S1A; Supplementary table 1**).

hPSC DE differentiation protocols consist of two steps that mimic the signals that sequentially specify DE from pluripotent epiblast cells ingressing through the anterior primitive streak (aPS) during gastrulation (Loh et al., 2014; Morgani and Hadjantonakis, 2020; Yiangou et al., 2018; Zorn and Wells, 2009). In the first step, Wnt and TGF-B/Nodal signaling promote exit from pluripotency and differentiation into aPS-like progenitors. In the second step, high TGF-B/Nodal signaling promotes the commitment of aPS progenitors into DE. This general approach has been applied to a variety of hPSC lines and can consistently generate >85% pure DE (CXCR4+, SOX17+ cells) (Loh et al., 2014). Based on previous work, we hypothesized that signaling requirements for mouse DE induction from EpiSCs should be generally similar to hPSCs, but that it was likely that the timing of each stage and the appropriate concentration of signaling factors would be different for mouse EpiSCs compared to hPSCs (Greber et al., 2010; Vallier et al., 2009)

To determine the optimal timing and concentration of each signal for aPS induction, we applied a gradient of increasing concentrations of CHIR99201 (to activate Wnt signaling via GSK3 inhibition) and ACTIVIN A (to activate TGF-B/nodal signaling) to DBA/2J EpiSCs for either 16, 20, or 24 hours. After aPS induction, all conditions were shifted to the same DE commitment conditions for an additional 24 hours, and DE induction purity (fraction of the total population co-expressing FOXA2+/SOX17+) was measured using immunostaining (**Fig. 2A, S1B**). Across all conditions, we observed DE purity ranging from 7–82%. The highest purity was achieved in conditions with higher ACTIVIN A and the highest levels of Wnt activation. Exposure to these aPS conditions for 16 hours was sufficient to enable DE commitment in the second stage of the protocol, and no further increase in purity was observed in the 20 or 24 hour aPS conditions.

Having defined these improved conditions for aPS induction, we next sought to define the optimal conditions for DE commitment of aPS-like cells (Stage 2) (**Fig. 2B, S1C**). To promote commitment of aPS-like cells into DE, DE differentiation protocols for hPSCs add a high concentration of ACTIVIN A for 48 hours. Concurrent inhibition of BMP and/or Wnt signaling has also been reported to reduce off-target differentiation of aPS cells into mesodermal lineages (Loh et al., 2014). Based on these data, we tested the effects of ACTIVIN A concentration, BMP inhibition (LDN-193189), Wnt inhibition (NVP-TNKS656, LGK-974), and the total duration of Stage 2 on DE commitment.

We induced aPS using the optimized conditions from **Fig. 2A** and examined the effects of different signaling conditions and timing on DE commitment (24, 36, 48 hours). We observed the highest purity of DE differentiation in conditions with maximal Activin A. Inhibition of BMP signaling markedly increased DE purity, whereas Wnt inhibition had minimal effects. DE purity was highest in the 24 hour Stage 2 condition, with longer durations leading to a substantial reduction of FOXA2/SOX17 positive DE cells. In the 24 hour condition, rare SOX17-negative cells occur in small clusters of densely packed cells. Most of these off-target cells co-express OCT4 and T, are FOXA2 (-), and lack expression of CDX2, which is expressed by paraxial mesoderm cells that are also generated from aPS-like progenitors. These data suggest that the $OCT4^{+}/T^{+}/FOXA2^{-}$ cells were most likely lagging aPS-like cells or nascent mesodermal cells. However, in the 36 and 48 hour conditions, SOX17 (-) cells could also result be properly specified DE cells that have started to adopt distinct A-P identities that no longer express SOX17, such as anterior or posterior foregut (Li et al., 2018).

Taken together, our systematic protocol optimization experiments demonstrate that the combination of Wnt activation in the presence of moderate TGF-B/Nodal signaling, followed by high TGF-B/Nodal activation and BMP inhibition can generate high purity DE from mouse EpiSCs in ~40 hours (**Fig. 2C, Supplementary Experimental Methods 1**).

To determine the purity of DE generated by our protocol more quantitatively, we used flow cytometry to measure the levels of the surface marker CXCR4 (expressed in endoderm and mesoderm) and the transcription factor SOX17 (expressed exclusively in DE) (**Fig. 2D**). Flow cytometry analysis indicated that the DE purity was ~95% using the optimized conditions. Next, we used this improved quantification method to assess the performance of our DE protocol on two additional EpiSC lines derived from distinct inbred strains of mice. The average DE purity achieved with a C57Bl/6J EpiSC line was similar to the DBA/2J line used for optimization (~95%), whereas the purity achieved

with an EpiSC line from the wild-derived inbred strain PWK/PhJ was slightly lower (~80%) (**Fig. 2E**). The DE purity was similar across technical replicates from independent experiments (n=3-4 technical replicates/line). It is important to note that these additional differentiation experiments were performed without any further line-specific optimization.

Taken together, these data show that mouse EpiSCs can be rapidly and efficiently differentiated into DE using conditions that are analogous to those currently used for hPSCs (Loh et al., 2014). However, we find that the major difference between hPSCs and EpiSCs is the timing required for each of the two steps in the DE differentiation protocol. EpiSCs differentiate more rapidly than hPSCs (40 vs. 72 hours), which is consistent with the faster pace of mouse development compared to human (Barry et al., 2017; Diaz-Cuadros et al., 2021; Matsuda et al., 2020; Rayon et al., 2020).

Molecular characterization of mouse EpiSC-derived definitive endoderm

To further characterize DE generated by our new protocol, we performed RNA-seq and ATAC-seq on DE generated from four independently derived C57Bl/6J EpiSC lines (characterized in **Fig. 1C-E**). CXCR4+ DE was purified with magnetic assisted cell sorting (MACS). EpiSC-derived DE exhibited large-scale transcriptional changes compared to EpiSCs, including upregulation of canonical DE marker genes (*Sox17*, *FoxA2*, *Cxcr4*, *Gata4/6*, *Cer1*) and down-regulation of pluripotency-associated genes, including *Oct4* (*Pou5f1*) and *Nanog* (**Fig. 3A, S2A-D**), consistent with results obtained using immunofluorescence staining (**Fig. S2E**) (Genga et al., 2019; Nowotschin et al., 2019). There are also widespread changes in chromatin accessibility at cis-regulatory elements in DE compared to EpiSCs. PCA of ATAC-seq data indicated that that cell type (EpiSC vs. DE) accounted for the majority of variation observed across samples (PC1; 92.5%) (**Fig. S2B**). Accordingly, we observed significant increases in chromatin accessibility at numerous putative enhancer elements near genes known to be critical for DE differentiation (*Foxa2, Gsc, Otx2, Sox17*) (**Fig. 3B**). Across the genome, cis-regulatory elements that gain chromatin accessibility in DE compared to EpiSCs are enriched for binding motifs for TFs known to play key roles in DE specification (GATA4/6, OTX2, SOX17, GSC). In contrast, cis-regulatory elements with higher chromatin accessibility in EpiSCs compared to DE exhibited enrichment for OCT4/POU5F1 and AP-1 (FOS/JUN) binding motifs (**Fig. 3C-E**). In summary, these genomic data demonstrate that DE exhibits global changes in gene expression and TF binding compared to EpiSCs and has the expected genomic features of *bona fide* PSC-derived DE.

Efficient and homogeneous conversion of EpiSCs into forebrain-patterned organoids

Given the success of our efforts to generate DE from EpiSCs, we next sought to determine whether primed ground state EpiSCs could also be a suitable starting cell type for generating neural organoids. Although the first neural organoid protocols were developed using embryoid bodies (EBs) formed from naïve mouse PSCs by Yoshiki Sasai and colleagues, these protocols were never widely adopted (Danjo et al., 2011; Eiraku and Sasai, 2011; Eiraku et al., 2011; Eiraku et al., 2008; Kadoshima et al., 2013; Nasu et al., 2012; Shiraishi et al., 2017; Wataya et al., 2008). We chose to focus our attempts on generating dorsal forebrain-patterned organoids, which have the potential to give rise to progenitors of multiple brain regions including the cerebral cortex, hippocampus, and thalamus (Montiel and Aboitiz, 2015).

Following extensive testing, we were able to identify conditions for robust EB formation, neural induction, and forebrain patterning of EpiSC-derived EBs (summarized in **Fig. 4A-B**). Briefly, EpiSCs are dissociated into single cells and EBs are formed from ~1,000 EpiSCs in neural induction media in Aggrewell plates. We found that addition of chroman 1 (ROCK inhibitor) and emricasan (pan-caspase inhibitor), components of the recently reported CEPT cocktail, help to dramatically reduce cell death during EB formation compared to commonly used ROCK inhibitors (Y-27632 or Thiazovivin; **Fig. S3A**) (Chen et al., 2021). To promote the generation of forebrain fates (telencephalon, diencephalon) we block Wnt and FGF signaling for the first two days of differentiation (Rifes et al., 2020; Takata et al., 2017; Tchieu et al., 2017). After the first 24 hours, EBs are removed from Aggrewell plates and embedded in Matrigel (based on the protocol developed by (Qian et al., 2018)). Under these conditions, many EpiSCs express SOX1 after the first 24 hours of EB formation, and nearly all cells in EBs express SOX1 after 48 hours (**Fig. 4C**). In contrast, mouse

neural organoid protocols that begin by generating EBs from naïve PSCs, have few SOX1+ cells after 72 hours in culture (Eiraku et al., 2008).

During early NE development, FGFs are posteriorizing, but rapidly switch to promoting telencephalic identity in the anterior-most region of the NE by ~E8.25 (Garel et al., 2003; Shimamura and Rubenstein, 1997). Therefore, we hypothesized that addition of FGF8 (FGF8b) after the first 48 hours of differentiation would facilitate induction of anterior telencephalic fates and could subsequently promote the proliferation and survival of FOXG1+ telencephalic neuroepithelial progenitors (Storm et al., 2006). By 96 hours (d4), EBs form continuous neuroepithelia, with consistent apicobasal polarity (**Fig. 4D-E, S3B**) (**Video S1**). Characterization of the complete 3D structure of d4 organoids using tissue clearing revealed mitotic neuroepithelial progenitors (phospho-histone H3+) at the apical membrane (marked by ZO-1), a central lumen, Nestin-expressing cells with characteristic radial-glial morphologies spanning the length of the developing tissue (apical-to-basal), as well as basal progenitors (TBR2+) and early-born postmitotic neurons migrating outward from the apical surface towards the basal lamina (**Fig. 4D-G**). We sparsely infected neuroepithelial progenitors with emGFP Sendai virus and used light sheet microscopy to live-image cell division in organoids from d3 to d4. This revealed radial glial cells undergoing interkinetic nuclear migration and symmetric division (Arai and Taverna, 2017; Noctor et al., 2004) (**Video S2)**. After d4, organoids are removed from Matrigel and switched into maintenance media (N2B27 media with vitamin A, and the neurotrophic factors BDNF and GDNF) and incubated in low adhesion cell culture dishes on an orbital shaker.

Next, we sought to characterize the specific regional identity of progenitors and neuronal populations generated in these organoids. At d4, all organoids expressed PAX6, which is expressed by progenitors in the caudal forebrain (diencephalon) and/or dorsal telencephalon, depending on the exact stage (Puelles et al., 2000). However, PAX6 expression was not homogeneous throughout the organoids, suggesting that not all neuroepithelial progenitors acquire the same regional identity under these conditions (**Fig. 5A**). By d8, most organoids spontaneously pattern into two opposing domains, possibly derived from the Pax6+ and Pax6- regions at d4 (Renner et al., 2017; Takata et al., 2017) (**Fig. 5B-G)**. Strikingly, one domain contained large numbers of PAX6+ post-mitotic neurons, which is a relatively specific feature of neurons derived from the prethalamus, which forms in the anterior-most domain of the diencephalon (prosomere 3, p3) (Caballero et al., 2014; Manuel et al., 2015) (**Fig. 5C).** In the domains opposite from the putative prethalamic domain, we observe neurons expressing markers of the cerebral cortex, such as TBR1, BCL11B/CTIP2, and BRN2, as well as RELN, a marker of Cajal-Retzius cells (**Fig. 5D-G, S4**). Taken together, these data suggest that organoids derived in these conditions polarize early into distinct domains that represent different anterior-posterior identities: the dorsal telencephalon and the anterior diencephalon (prosomere 3). This phenomenon was also observed in independent experiments using a second EpiSC line (C57Bl/6J) (**Fig. S4C-D**).

Characterization of forebrain organoids using scRNA-seq

To gain further insight into the precise cell populations present within forebrain organoids, we generated scRNA-seq data from forebrain organoids at d12 (5,652 total cells). Reduced-dimensionality visualization using UMAP revealed multiple distinct clusters, which we tentatively assigned cell identities based on their expression of known marker genes. Importantly, nearly all cells within the organoid appear to be neuroectodermal derivatives, suggesting that neural induction in EBs is highly efficient (**Fig. 6, S5A, Supplementary table 5**)

Consistent with our immunostaining results, we observed clusters that appeared to be immature (cluster 7) and more mature (cluster 8) prethalamic neurons. The prethalamic domain generates multiple distinct types of GABAergic neurons that form the thalamic reticular nucleus and the zona incerta (Puelles et al., 2013). The early pre-thalamic neuron cluster (cluster 7) was characterized by expression of markers associated with early development of GABAergic neurons (*Ascl1*, *Dlx5*, *Dlx6, Gad1, Gad2*), which is consistent with the fact that most prethalamic neurons are GABAergic. The more mature pre-thalamic neuron cluster (Cluster 8) was characterized by expression of the prethalamic neuron markers (*Pax6*, *Isl1*, *Meis2*, *Mmp16)* (Mandai et al., 2014; Nagalski et al., 2016; Ono et al., 2014; Virolainen et al., 2012) (**Fig. 6E,D**). Among the more mature prethalamic neurons, there was evidence of further distinction based on the relative expression levels of *Ptprd* and *Pax6*, which distinguishes between the lateral or the medial portions of the developing prethalamus *in vivo* (Guo and Li, 2019; Kim et al., 2020; Li et al., 2020; Sommer et

al., 1997; Tuttle et al., 1999). In addition, more mature prethalamic neurons express higher levels of *L1cam* and *Nrcam*, which play a role in the pathfinding of axons from prethalamic neurons that project into the thalamus (Molnar et al., 2012; Quintana-Urzainqui et al., 2020) (**Fig. S5B**).

We identified one cluster (Cluster 6) that appeared to correspond to the glutamatergic neurons expressing TBR1 and/or BCL11B/CTIP2 that we observed identified by immunostaining. However, although expression of TBR1/CTIP2 is often associated with cerebral cortex development, neurons in this cluster instead express several markers specific to hippocampal neurons (*Zbtb20*, *Neurod1*, *Prox1),* and lack expression of the telencephalic marker *Foxg1*, which is expressed broadly in the cerebral cortex but absent from the hippocampus (Hatami et al., 2018) (**Fig. 6F)**. The lack of FOXG1 was further confirmed by immunofluorescence staining (data not shown). These data suggest that these neurons are more likely to be developing hippocampal neurons than deep layer cortical neurons. The observed similarity in marker profiles between these two cell types is consistent with *in vivo* data showing that the developing neocortex (cerebral cortex) and allocortex (hippocampus and olfactory bulb) have similar expression profiles during embryonic development (Cadwell et al., 2019).

We also observed a cluster of cells (Cluster 3) that resemble cortical hem, and are characterized by expression of *Rspo* genes, *Wnt8b*, *Wnt9a,* and *Bmp7* (Grove et al., 1998; Hasenpusch-Theil et al., 2012) (**Supplementary table 5**). The cortical hem is a secondary organizer located next to the developing hippocampus in the dorsal-most region of the developing telencephalon (Vieira et al., 2010). In addition to its role in cortical patterning via secretion of Wnt and BMP signals, neural progenitors in the cortical hem also generate Cajal-Retzius cells during early stages of telencephalon development (Takiguchi-Hayashi et al., 2004). Accordingly, we observe cells bridging clusters 3 and 6 that co-express the Cajal-Retzius cell markers *Trp73* (p73) and *Reln (reelin)* (Hanashima et al., 2007; Simon et al., 2012) (**Fig. 6F**).

We also used VoxHunt (Fleck et al., 2021) to compare our single cell dataset to *in situ* hybridization data from the E18 mouse brain (Allen Developing Mouse Brain Atlas) (Thompson et al., 2014). This analysis confirmed that all major cell clusters in our data most closely resemble medial pallium (hippocampus/cortical hem) and prosomere 3 (p3, prethalamus) (**Fig. 6H**).

In addition to these neuronal clusters, we also found clusters (Cluster 4 and 5) that appear to be oligodendrocyte progenitor cells (OPCs), characterized by high expression of *Olig1*, *Olig2*, *Pdgfra*, and *Dll1* (**Fig. 6G**, **Supplementary table 5**). Among all of the clusters, OPCs were the only cell type with a significant population of actively proliferating cells at d12 (**Fig. 6D**). Immunofluorescence staining for OLIG2+ at d12 confirmed the presence of these cells, which were intermingled with CTIP2+ hippocampal neurons, suggesting that the same progenitors that give rise to the hippocampal glutamatergic neurons progress to produce OPCs (**Fig. S5C**). This is consistent with prior studies showing that mouse radial glial cells from the dorsal telencephalon give rise to OPCs after birth, and explant studies demonstrating that dorsal telencephalic progenitors from the embryonic brain often generate OPCs when explanted or cultured *in vitro* (Kessaris et al., 2006; Qian et al., 1997; Tekki-Kessaris et al., 2001).

Taken together, these data reveal that our EpiSC-based forebrain organoid protocol can robustly generate organoids containing multiple distinct regions of the developing forebrain, including dorsal, caudo-medial telencephalon (hippocampus and cortical hem) and dorsal anterior diencephalon (prosomere 3, prethalamus). These two domains are adjacent to one another, each bordering the boundary between the telencephalon and diencephalon. Prethalamic domains and hippocampal/cortical hem domains generally formed on opposite poles of organoids, suggesting that organoids undergo spontaneous polarization during early stages of develop and develop into distinct domains representing different anterior-posterior identities (Takata et al., 2017).

DISCUSSION

Directed differentiation models enable a range of experimental techniques that would be difficult or impossible to perform *in vivo*, including genetic and chemical screens, live imaging, and experimental embryology approaches such as *in vitro* reconstitution (Hadjantonakis et al., 2020; Li et al., 2019; Matsuda et al., 2020; Moris et al., 2020; Rayon et al., 2020; Schlissel and Li, 2020). Using mouse PSCs for directed differentiation has multiple advantages compared to

human PSCs. For example, using mouse PSCs makes it possible to use the extensive genetic tools and resources that have been generated for mice. This includes not only knockout lines, reporter lines, and human disease models, but also numerous distinct inbred strains and outbred heterogeneous stocks such as Diversity Outbred mice that can be used for quantitative genetic approaches (Churchill et al., 2012; Lager et al., 2018; Lusis et al., 2016; Skelly et al., 2020). In addition, results from mouse PSC differentiation models can be directly compared to genetically-matched tissues from mouse embryos, which allows for rigorous assessments of model validity (Voelkl et al., 2020). Such data represent an important proof-of-concept for hPSC-based cellular models of development and disease processes, as it will never be possible to perform similar cross validation experiments in developing human embryos (Kleiman and Engle, 2021).

Despite the advantages of mouse directed differentiation, mouse PSC models are much less frequently used than human PSC-based models. This is likely due in part to the limited availability of mouse PSC directed differentiation protocols that are robust, efficient, generalizable, and user-friendly. We found that by using mouse ground state EpiSCs, it is possible to develop protocols for directed differentiation as robust as state-of-the-art hPSC protocols. Since naïve PSCs can be easily converted into ground state EpiSCs *in vitro* within a few days, our new protocols can be applied to generate DE and neural organoids from pre-existing naïve PSC lines that carry specific genetic modifications, including gene knockouts or fluorescent reporters. These new optimized methods for directed differentiation represent a valuable addition to the experimental toolkit available for studying mouse development and thus should be of substantial interest to the developmental and stem cell biology research communities.

Rapid and efficient generation of definitive endoderm from EpiSCs

By adapting protocols developed for hPSCs, we demonstrate that ground state EpiSCs can be differentiated into definitive endoderm with high purity (>90%) within 40 hours. Our findings are consistent with previous work showing that EpiSCs and hPSCs respond similarly to most differentiation cues (Greber et al., 2010; Vallier et al., 2009). The most significant difference we observe between hPSC and EpiSC differentiation is timing. The optimal timing for EpiSCs is 16 hours for aPS induction and 24 hours to DE commitment, whereas hPSC protocols use 24 hours for aPS induction and 48 hours to DE commitment (Loh et al., 2014). These differences are consistent with the faster pace of mouse development compared to human, which has been previously observed when directly comparing mouse and human differentiation protocols (Ebisuya and Briscoe, 2018).

The high purity DE generated by our protocol will be an ideal starting point for subsequent differentiation into progenitors of a variety of endodermal organs, such as lung, liver, or intestinal tissues or organoids (Fowler et al., 2020; Yiangou et al., 2018). The high efficiency of this protocol can also facilitate large-scale genetic screens to identify genes that control endoderm commitment (Li et al., 2019). It will be interesting to compare the transcriptional regulatory networks and cis-regulatory elements controlling mouse and human endoderm differentiation in future studies.

Improved methods for mouse neural organoid generation

The first neural organoid protocols were developed using naïve mouse PSCs by Yoshiki Sasai and colleagues (Danjo et al., 2011; Eiraku and Sasai, 2011; Eiraku et al., 2011; Eiraku et al., 2008; Kadoshima et al., 2013; Kelava and Lancaster, 2016; Nasu et al., 2012; Shiraishi et al., 2017; Wataya et al., 2008). Although these pioneering studies provided the foundation for the subsequent development of protocols to generate neural organoids from hPSCs, the original mouse protocols were never widely adopted and are now rarely used. This is due, at least in part, to the fact that these protocols were less robust and user-friendly than current state-of-the-art hPSC organoid protocols, and thus they proved too difficult for other labs to readily adopt. In addition, most of these early mouse neural organoid protocols pre-date the development of improved methods for consistently generating polarized neuroepithelia in neural organoids (e.g. matrigel embedding or addition of laminin/matrigel to the culture media) (Lancaster et al., 2013; Nasu et al., 2012). Therefore, we reasoned that by starting with EpiSCs instead of naïve PSCs, and by implementing improved techniques for organoid generation developed for hPSC-derived organoids, we would be able to develop an improved protocol for the generation of mouse neural organoids.

Although our initial goal was to generate a mouse organoid model of cerebral cortex development, we instead observed domains of hippocampal and prethalamic progenitors. This suggests that our conditions for anterior-posterior patterning must be further modified to generate cerebral cortex, which comes from the most anterior region of the developing neuroectoderm. Previous work with mouse neural organoids found that insulin exerts posteriorizing effects on developing neuroectoderm via activation of FGF signaling, and it is necessary to remove insulin from the media at early stages of neural induction to maintain anterior neuroectoderm patterning (Takata et al., 2017; Wataya et al., 2008). However, this strategy is not readily adaptable to primed PSC organoid models, because insulin is critical for primed pluripotent stem cell growth and survival, making it difficult to form EBs and perform neural induction in insulin free conditions (as was done for naïve PSCs). Based on these studies, we added FGFR inhibitors in addition to Wnt inhibitors for the first two days to block the posteriorizing influence of FGF signaling. However, we still observed spontaneous polarization of developing organoids into telencephalic and diencephalic regions (Renner et al., 2017; Takata et al., 2017). Interestingly, for hPSC-based cerebral cortex organoids, inhibition of Wnt signaling (in addition to dual SMAD inhibition for neural induction) is generally sufficient to maintain anterior neuroectodermal identity and specification of cortical fates, even in the presence of high levels of insulin. It will be interesting to explore the mechanisms underlying this apparent difference in anterior-posterior patterning between mouse and human.

Methods for the generation of organoids containing diencephalic progenitors have previously been reported (Shiraishi et al., 2017; Xiang et al., 2019). These protocols use a combination of moderate insulin, inhibition of FGF/ERK, and activation of BMP signaling (BMP7) to specify diencephalic progenitors of the thalamus and epithalamus (prosomere 2) and pretectum (prosomere 1). In contrast, we find that addition of high levels of FGF8b and no exogenous BMP7 induces prethalamic (prosomere 3) fates. This is consistent with the high expression of FGF8b in the dorsal midline of p3 and prethalamic eminence at early stages of diencephalon development (E10.5-E12.5) (Martinez-Ferre and Martinez, 2009). *In vivo*, prethalamic progenitors generate numerous distinct types of inhibitory (GABAergic) neurons that form the thalamic reticular nucleus and the zona incerta (Li et al., 2020). Prethalamic neurons have not been generated efficiently in previous mouse and human thalamic organoids, and thus our forebrain organoid protocol should be useful for studying the development of this complex and poorly understood brain region in greater depth.

Mouse forebrain-patterned organoids can complement *in vivo* mouse models of forebrain development as well as neurodevelopmental disorders (Sestan and State, 2018). Our neural organoid model should facilitate direct comparisons between the phenotypes observed *in vivo* in mice with mutations associated with neurodevelopmental disorders, and those observed in analogous mouse organoid models *in vitro*. In contrast, hPSC-based models of neurodevelopmental disorders cannot be conclusively validated *in vivo*, and thus it is difficult to evaluate the relevance of observations made in neural organoid models *in vitro* for developmental pathophysiology *in vivo*. More rigorous assessment of the strengths and limitations of *in vitro* models in mice can help to inform experimental design for hPSC-based modeling of neurodevelopmental disorders. In future studies, it will be interesting to perform in depth comparisons between developing neural organoids and their analogous cell types *in vivo*, and to compare mouse and human organoid development.

Limitations of the study

Although we were able to develop robust protocols for differentiation of ground state EpiSCs into DE and neural organoids, it is likely that ground state EpiSCs will not be an ideal starting point for differentiation into all cell types. For some protocols, epiblast-like cells (EpiLCs) or more recently described stable formative stage stem cells might be superior to ground state EpiSCs (Hayashi et al., 2011; Kinoshita et al., 2020; Morgani et al., 2018; Yu et al., 2021). Even when cultured with Wnt inhibition, we observe some heterogeneity in EpiSC cultures, marked by expression of Foxa2/T (Bernemann et al., 2011; Blauwkamp et al., 2012; Song et al., 2016; Tsakiridis et al., 2014). Additional optimization of ground state EpiSC conditions could help to develop conditions that are as robust as human PSCs and further mitigate issues with current culture conditions such as residual heterogeneity. The impact of long-term culture on genetic and epigenetic stability of ground state EpiSCs has not been characterized in depth (Halliwell et al., 2020). Similarly, the long-term effects of Tankyrase inhibition on the genetic and epigenetic stability of EpiSCs remain to be characterized. Several types of Wnt inhibitors have been used for EpiSC culture, and it remains uncertain how different

means of inhibiting Wnt activity might change the properties of ground state EpiSCs (Kurek et al., 2015; Sumi et al., 2013; Wu et al., 2015). There is an emerging consensus in the hPSC directed differentiation field that the initial culture conditions of hPSCs have a significant impact on their ability to differentiate into specific lineages (Cornacchia et al., 2019; Watanabe et al., 2019). Thus, it is likely that further improvements of ground state EpiSC culture will be the best way to generally improve their performance in directed differentiation experiments.

MATERIALS AND METHODS

Summary

Step-by-step protocols for directed differentiation experiments are provided as **Supplementary Experimental Methods 1**

For information about antibodies, see **Supplementary Table 2**

For reagent information, see **Supplementary Table 3**.

Detailed information for each cell line is provided in **Supplementary Table 4.**

Complete UMAPs, related to figure 6, are provided in Supplementary Table 5.

Mouse ESC culture and conversion to EpiSCs

All mouse ESC lines were maintained on gelatin-coated dishes with irradiated mouse embryonic fibroblast feeder cells using serum/LIF media comprised of DMEM (high glucose, GlutaMAX, HEPES), 1% nonessential amino acids, 1% sodium pyruvate, 1% penicillin-streptomycin, 0.1% 2-mercaptoethanol, 10% fetal bovine serum and 1000 units/mL ESGRO LIF. All ESC lines (except lines from C57B/6J X 129 F_1 hybrids) were cultured in serum/LIF media with 2i containing 3 uM CHIR99201 and 1 uM PD0325901. Media was changed daily, and ESCs were passaged upon 70% confluence at a 1:6 ratio using TrypLE. ESC-to-EpiLC conversion was performed as described in (Morgani et al., 2018). Briefly, ESCs were lifted using 0.1 u/uL Collagenase IV, centrifuged and dissociated into a single cell solution using Accutase. The ESC suspension was plated on Fibronectin coated plates (16.7 ug/mL) at a density of 17,500 cells/cm^2 in N2B27 media supplemented with 12.5 ng/ml heat stable recombinant human bFGF, 20 ng/ml Activin A, and 1% Knockout Serum Replacement. N2B27 media consists of 50% DMEM-F12, 50% Neurobasal, 0.5% N2 supplement, 1% B27 supplement without vitamin A, 2 mM glutamax, 1% penicillin-streptomycin, and 0.1% 2-mercaptoethanol. The media was changed after 24 hours, and cells were converted to EpiSCs after 48 hours. For EpiLC-to-EpiSC conversion, EpiLCs were dissociated into small clumps (~3-5 cells) with Accutase and plated at a density of ~50,000-100,000 cells per cm^2 on irradiated mouse fibroblast feeder cells in N2B27 supplemented with 20 ng/ml activin A, 12.5 ng/ml heat stable bFGF and Wnt inhibitor (175 nM NVP-TNKS656). EpiSC media was changed daily, and cells were passaged every ~48 hours at a 1:6 ratio using 0.1 u/ul collagenase IV followed by dissociation with Accutase into small clumps of 3-5 cells.

Definitive endoderm differentiation

For definitive endoderm differentiation, EpiSCs were generally used after ~4-8 passages in EpiSC conditions. For differentiations, high quality cultures of EpiSCs were detached using Collagenase IV (0.1 u/µL), washed once with PBS, and then dissociated into a single cell suspension using Accutase. EpiSCs were plated at a density of 110,000 cells/cm^2 in chemically defined media (CDM) containing 50% IMDM, 50% Ham's F12 Nutrient Mix with Glutamax, 1% chemically defined lipid concentrate, 450 uM Monothioglycerol, 1% polyvinyl alcohol (w/v), 15 ug/ml Apo-transferrin, 0.5% Glutamax, 0.7 ug/ml Insulin and supplemented with 20 ng/ml Activin A, 12.5 ng/ml heat stable bFGF, 175 nM NVP-TNKS656, 1% knockout serum replacement, and 2 uM Thiazovivin. Plating media was removed 6 hours after seeding, cells were washed gently with PBS without calcium or magnesium (PBS-/-) and the media was changed to CDM supplemented with 3 uM CHIR99201 and 40 ng/ml Activin A. 16 hours after the first media change, cells were washed with PBS-/- and the media was replaced with CDM supplemented with 100 ng/ml Activin A and 100 nM LDN-

193189. After 24 hours (46 total hours after cells were initially seeded), cells were fixed for immunostaining or collected for further analysis.

Immunostaining

For immunostaining of adherent cells (EpiSCs and DE), cells were rinsed twice with PBS-/- and fixed with 4% paraformaldehyde solution for 30 min at room temperature. After fixing, cells were washed twice with PBS -/- and permeabilized with PBS -/- containing 1% Triton-X (PBST) for 10 minutes at room temperature and blocked for 30 minutes with PBS -/- containing 0.3% Triton-X and 5% fetal bovine serum. Primary antibodies were diluted in PBST with 1% fetal bovine serum and incubated overnight at 4C. Detailed information about the antibodies used can be found in **Supplementary Table 2**. Following overnight incubation and three 10 min washes with PBST, cells were incubated with secondary antibodies for 2 h at room temperature. Cell were rinsed with two 10 min washes of PBST before DAPI staining (1 μM) for 15 min at room temperature. Cells were rinsed with PBST and imaged using a Leica Dmi8.

Quantification of Immunofluorescence images

All image quantification was performed using Cell Profiler version 4.2.1 to calculate efficiency of DE differentiations via FOXA2 and SOX17 expression. DAPI-stained nuclei were identified using the Identify Objects module. To identify nuclei stained with each antibody, segmented nuclei were further analyzed using the Identify Objects module and the Relate Objects Module. These data were then classified as either positive or negative for each marker using the Filter Objects module and overall efficiency was calculated by percentage of expression for each marker.

Flow Cytometry

Following DE differentiation, cells were dissociated in Accutase and washed in FACS buffer consisting of PBS-/-, 20% fetal bovine serum and 0.1% 0.5 M EDTA. CXCR4 antibody conjugated to Alex647 and live/dead zombie UV staining were diluted in FACS buffer at the manufacturer's recommended concentration and incubated for 15 minutes at room temperature. Cells were washed and resuspended in FACS buffer with fixation/permeabilization diluent and fixation/permeabilization concentrate for 30 minutes at room temperature. The cells were then resuspended in permeabilization buffer with SOX17 antibody conjugated to Alexa488 at the manufacturer's recommended concentration for additional 30 minutes at room temperature. Fixed cells were then analyzed using the Cytek Aurora. Flow cytometry results were analyzed using FlowJo software v10.8 and single color stainings were used as a negative control for gating purposes.

MACS purification of Definitive Endoderm

Definitive endoderm cells were isolated for ATAC-seq and RNA-seq analyses using MACS following manufacturer's instructions, with slight modifications. Cells were dissociated in Accutase, washed in FACS buffer, counted, and resuspended with Rat anti-mouse CD184/CXCR4 Alexa Fluor 647 conjugated antibody at a concentration of 1 uL/10^6 cells and incubated for 1 h at 4°C. Cells were then washed with FACS buffer, centrifuged, and resuspended in a solution composed by 80% FACS buffer and 20% Anti-Rat IgG MicroBeads. Cells were incubated in the mixture for 15 min at 4°C, then washed and resuspended in 500 uL buffer. The MS column was placed in OctoMACS separator on the MACS multistand. The cell suspension was added to the column and rinsed 3 times with buffer. Finally, the column was removed from separator and the magnetically labeled cells were collected.

RNA extraction for RNA-seq

Phase separation in cells lysed in TRIzol was induced with 50% isopropanol + 0.5% 2-Mercaptoethanol and RNA was extracted from the aqueous phase using the MagMAX mirVana Total RNA Isolation Kit on the KingFisher Flex Magnetic Particle Processor according to the manufacturer's protocol with 1-2M cells input. Samples were eluted in 38 μL elution buffer.

Transcriptome sequencing of EpiSC samples

After RiboGreen quantification and quality control by Agilent BioAnalyzer, 100-500 ng of total RNA with RIN values of 9.5-10 underwent polyA selection and TruSeq library preparation according to instructions provided by Illumina (TruSeq Stranded mRNA LT Kit), with 8 cycles of PCR. Samples were barcoded and run on a NovaSeq 6000 in a PE100

run, using the NovaSeq 6000 S4 Reagent Kit (200 Cycles) (Illumina). An average of 36 million paired reads were generated per sample and the percent of mRNA bases per sample ranged from 84% to 89%.

Transcriptome sequencing of Definitive Endoderm samples

After RiboGreen quantification and quality control by Agilent BioAnalyzer, 2 ng total RNA with RNA integrity numbers ranging from 7.7 to 10 underwent amplification using the SMART-Seq v4 Ultra Low Input RNA Kit, with 12 cycles of amplification. Subsequently, 10 ng of amplified cDNA was used to prepare libraries with the KAPA Hyper Prep Kit using 8 cycles of PCR. Samples were barcoded and run on a NovaSeq 6000 in a PE100 run, using the NovaSeq 6000 S4 Reagent Kit (200 Cycles) (Illumina). An average of 39 million paired reads were generated per sample and the percent of mRNA bases per sample ranged from 83% to 87%.

ATAC-seq

ATAC-seq was performed as previously described (Buenrostro et al., 2015) and (dx.doi.org/10.17504/protocols.io.bv9mn946) with minor modifications. Briefly, cells were dissociated into single cells, filtered through a Flowmi cell strainer and the nuclei were isolated by incubation with lysis buffer (10 mM Tris-HCl pH7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% Tween-20, 0.1% NP-40, 0.01% digitonin and 1% BSA) for 5min at 4C. 20,000 and 40,000 nuclei were isolated, and the DNA was tagmented with Tn5 (Illumina) and amplified using NEBNext® High-Fidelity 2X PCR Master Mix (NEB). The number of cycles was estimated by qPCR. DNA tagmentation efficacy was evaluated with bioanalyzer 2100 (Agilent technologies) and the DNA amounts calculated with Qubit. i5 and i7 primer sequences were obtained from (Mezger et al., 2018). The resulting DNA libraries were sequenced using the NextSeq550 system (Illumina) and about 25 million reads were obtained per sample in duplicate.

RNA-seq Analysis

Resulting fastq files were mapped to their transcripts and quantified using salmon (Patro et al., 2017) against the genome version mm10. For downstream analyses, Rstudio was used with the packages Deseq2, PCAExplorer and pheatmap (Love et al., 2014; Marini and Binder, 2019). Statistical significance was calculated using a Wald test and corrected using the Benjamini-Hochberg method.

ATAC-seq Analysis

The raw fast data were analyzed using Basepair software (https://www.basepairtech.com/) with a pipeline that included the following steps. The raw reads were trimmed using fastp to remove low-quality bases from reads (quality < 20) and adapter sequences. The trimmed reads were aligned using Bowtie2 (Langmead and Salzberg, 2012) to UCSC genome assembly mm10. Duplicate reads were removed using Sambamba. Peaks were identified with MACS2 (Gaspar, 2018) and those overlapping with satellite repeat regions were discarded. For further analyses, a union peak atlas was created from the MACS2 files. Peak intensity for each sample was counted using featureCounts (Liao et al., 2014). HOMER (v4.11) (Heinz et al., 2010) was used for motif analyses. Principal component analyses were done using the R packages DESeq2 (Love et al., 2014) and PCAexplorer (Marini and Binder, 2019).

For motif enrichment analysi, ATAC-seq peaks in the atlas were associated with TF motifs in the CIS-BP database (Weirauch et al., 2014) using FIMO (Grant et al., 2011) of MEME suite (Bailey et al., 2009), under the P value cutoff of 1e-4. We limited the motif analyses to motifs for 236 TFs that are expressed in EpiSCs or DE (RPKM > 5). Normalized counts were obtained using variance-stabilizing transformation in DEseq2. The shift in the cumulative distribution of normalized counts was compared between the subset of the atlas containing the TF motif and the total atlas using one-sided Kolmogorov-Smirnov test.

Generation of Forebrain Organoids

For forebrain organoid generation, B6129SF1/J EpiSC colonies were detached using Collagenase IV (0.1 u/μL), washed once with $PBS^{-/-}$, and then dissociated into a single cell suspension using Accutase. EpiSCs were then seeded at 1,000 cells/microwell in an Aggrewell plate in EB formation media containing 50 nM chroman-1, 5 μM emricasan, 100 nM LDN, 10 μM SB431542, 100 nM PD173074 and 100 nM LGK974 in N2B27 (B27 without vitamin A) media. After 24h, the EBs were recovered and transferred to the same media but without chroman-1 and emricasan and embedded in Matrigel as previously described (Qian et al., 2018). Briefly, 66.7 μL of EBs + media were mixed with

100uL of Matrigel using wide-bore tips. The mix was added to a non-adherent 6-well plate without touching the edges of the well and incubated at 37C for 30min. After 30 min, 3mL of warm media were added on top. At day 2, the media was changed to plain N2B27 (B27 without vitamin A) with 100 ng/mL Fgf8b and kept for 48 h more. Day 4 EBs were recovered from Matrigel using cell recovery solution (Corning) and incubating for 30 min at 4C. After day the EBs were transferred to a Petri dish and kept in N2B27 (B27 Supplement with vitamin A) with BDNF and GDNF while shaking at 65 rpm (Infors HT, Celltron benchtop shaker). For sparse labeling of neuronal progenitor cells with CytoTune EmGFP Sendai fluorescence reporter, 7.5 µL were added to the 2 mL media in the Aggrewell plate at D0 and kept until D1. The organoids were then imaged for ~24 hours from d3 to d4.

Immunofluorescence Staining of Organoids

For organoid staining, the organoids were collected in tubes pre-coated with 1%BSA. The samples were washed twice with $PBS^{-/-}$ and fixed with 4% PFA for 45 min at 4C. Afterwards, they were washed twice with PBS-/- and permeabilized with 0.5% Triton-X from 15 min at 4C. Blocking was performed for 15 min in organoid washing solution (OWS, adapted from (Dekkers et al., 2019), consisting of 0.2% Triton-X, 0.02% SDS and 0.2% BSA in PBS-/-) at 4C. Afterwards, the samples were incubated in OWS with primary antibodies overnight at 4C. On the following day they were washed 3x for 2h each with OWB while shaking, to then be incubated overnight at 4C with secondary antibodies and DAPI (1 µM) in OWB. On the following day, the organoids were once again washed 3x with OWB for 6 h in total while shaking. After that, the organoids were mounted in DeepClear solution (Celexplorer). The samples were imaged using the confocal Nikon A1RHD25 or Leica SP8 and analyzed with the Imaris software. For live imaging, the organoids were embedded in phenol-free Matrigel and imaged using the TrueLive3D Imager system (Luxendo).

Single-cell RNA sequencing

Multiple organoids were collected and mixed at day 12, washed with PBS-/- and dissociated with Accutase until a single cell suspension was achieved. Then, the sample was washed once again with PBS-/-. The resulting cell suspension was used for 10X scRNA-seq (10x Genomics, Single Cell 3' Kit v3.1, dual index) following the manufacturer's directions. 326M reads were obtained. The resulting fastq files were processed with CellRanger (10X genomics cloud). CellBender was used to eliminate background reads and other artifacts (Fleming et al., 2019), additionally, any cell with more than 10% mitochondrial reads was excluded. The Seurat package (v4) was used for downstream analyses (Hao et al., 2021). VoxHunt analyses were performed as described in (Fleck et al., 2021).

ACKNOWLEDGEMENTS

This study would not have been possible without the core facilities at MSKCC. We would like to specifically thank Dr. Yas Furuta (Mouse Genetics Core facility, MSKCC), Murray Tipping and Vitaly Boyko (Molecular Cytology Core, MSKCC), Dr. Richard Koche (Epigenetics Innovation Lab, MSKCC), Dr. James Muller (Developmental Biology Imaging Core, MSKCC), Dr. Ronan Chaligne (Single Cell Research Initiative, MSKCC), and members of the Integrated Genomics Operation Core (MSKCC) for their advice and assistance with experiments and data analysis throughout the course of this project.

We would like to thank Dr. Laura Reinholdt and Anne Czechanski (Jackson Labs) for sharing PWK/PhJ embryonic stem cell lines, Dr. Christopher Baker and Dr. Candice Byers (Jackson Labs) for sharing C57Bl/6J embryonic stem cell lines, and Dr. Matthias Stadtfeld (Weill Cornell Medical College) for sharing B6129SF1/J embryonic stem cell lines.

We would like to thank our colleagues Dr. Anna Katerina Hadjantonakis, Dr. Danwei Huangfu, Dr. Lorenz Studer, Dr. Nan Yang, Dr. Matthias Stadtfeld, Dr. Eftychia Apostolou, Dr. Sophie Morgani, Dr. Clayton Schwarz, Dr. Gabriele Ciceri, Dr. Ryan Walsh, Bess Rosen (Huangfu Lab), Renhe Luo (Huangfu lab), Dapeng Yang (Huangfu Lab) for helpful discussions and advice throughout the course of the study. Figure panels 3C and 4A were created using Biorender.com

We would like to thank Dr. Nan Yang (Mount Sinai), Dr. Matthias Stadtfeld, Dr. Timothy Cherry for helpful comments on the manuscript.

This study was made possible by financial support from: Startup funding from the Sloan Kettering Institute and Memorial Sloan Kettering Cancer Center (T.V.) and the Josie Robertson Investigator Program (T.V.), NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation (T.V.)., Tri-Institutional Stem Cell Initiative (Starr Foundation), and a training award from NYSTEM (contract # C32559GG) and the Center for Stem Cell Biology at MSK (D.M.).

AUTHOR CONTRIBUTIONS

Conceptualization: T.V., D.M.C., E.K.C.
Methodology: D.M.C., E.K.C., R.A.G., T.V.
Investigation: D.M.C., E.K.C, R.A.G., M.I., Y.L., H.C., J.K., T.V.
Data curation: D.M.C.
Writing: D.M.C., E.K.C., R.A.G., T.V.
Funding Acquisition: T.V.
Supervision: T.V.

DATA AVAILABILITY

Source data and processed files are available in the Gene expression Omnibus under the accession numbers GSE189869 (RNA and ATAC-seq datasets) and GSE189870 (scRNA-seq).

DECLARATION OF INTERESTS

The authors have no competing interests to declare.

FIGURE LEGENDS

Figure 1: Derivation and characterization of mouse EpiSCs cultured in primed ground state conditions. (**A**) Phase contrast image of an EpiSC colony. Scale bar = 50μm. (**B**) Immunofluorescence images of EpiSC cultures. Markers of general pluripotency (SOX2, OCT4), primed pluripotency (OTX2), naive pluripotency (KFL4, NANOG), and lineage commitment (T, FOXA2, SOX1). Scale bar = 50 μm. (**C**) Gene expression heatmap of RNA-sequencing data from EpiSC lines (n=12) for selected genes associated with different stages of pluripotency and lineage-commitment. Each EpiSC line was derived from a distinct mouse ESC line via *in vitro* naïve-to-primed conversion, and represent biological replicates. (**D**) Principal component analysis (PCA) plot of RNA-seq data (n =12 biological replicates; C57Bl/6J = 5, B6.129 = 3, PWK/PhJ = 3, DBA/2J = 1) and (**E**) ATAC-seq data (n =9 biological replicates; C57Bl/6J = 3, B6.129 = 2, PWK/PhJ = 3, DBA/2J = 1) for EpiSC lines. (**F**) PCA plot of ATAC-seq data for EpiSC lines (n =7 biological replicates; C57Bl/6J = 4, PWK/PhJ = 2, DBA/2J = 1). For comparison, ATAC-seq data from an additional set of EpiSC lines EpiSC-AF (cultured with Activin A and Fgf2) and EpiSC-AFX (cultured with Activin A, Fgf2, and Wnt inhibitor XAV939) from (Kinoshita et al., 2020) and ATAC-seq data from several early embryonic lineages *in vivo* (E6.5 epiblast, E6.5 visceral endoderm, and E7.5 ectoderm, E7.5 mesoderm, E7.5 endoderm, and E7.5 primitive streak were included from (Xiang et al., 2020).

Figure 2: Systematic optimization of conditions for EpiSC-DE differentiation. (**A-B**) Results of DE protocol optimization experiments. For each condition, immunofluorescence staining was performed following 40 h of differentiation (end of Stage 2). Immunofluorescence images for the indicated antibodies were quantified using CellProfiler. Markers of DE (FOXA2, SOX17), pluripotency (OCT4), and PS/mesoderm (T, CDX2) were examined. (**A**) Summary of results from anterior primitive streak (Stage 1) optimization experiments (n=2 technical replicates). (**B**) Summary of results for optimization of DE commitment (Stage 2) conditions (n=2 technical replicates). For Stage 1, the optimal conditions identified in (**A**) were used, and then cells were switched into the indicated conditions for Stage 2. (**C**) Diagram of the optimized protocol for differentiation of mouse EpiSCs into DE and a current human PSC to DE protocol (from Loh et al, 2014). (**D**) FACS Quantification of immunostaining for DE markers (SOX17-Alexa488 and CXCR4-Alexa647) from cells at the endpoint of the optimized DE differentiation protocol. (**E**) Quantification of DE differentiation purity (SOX17/CXCR4 double positive cells) across multiple experiments performed using the new optimized DE differentiation protocol. Each dot indicates a technical replicate from the indicated line (C57Bl/6J, n=4; DBA/2J, n=3; PWK/PhJ, n=3).

Figure 3: Genomic characterization of EpiSC-derived definitive endoderm. (**A**) RNA-sequencing data comparing gene expression between C57Bl/6J EpiSCs (n=5, biological replicates) and C57Bl/6J MACS-sorted DE (n=4, biological replicates). Genes exhibiting significant differences in expression between EpiSCs and DE are highlighted in orange ($P < 0.01$, Wald test and corrected using the Benjamini-Hochberg method). A set of critical genes for EpiSCs and DE are further annotated in red. (**B**) Chromatin accessibility (ATAC-seq) data from EpiSC (top row) or DE (bottom row) samples. Loci of genes known to play critical roles in DE development are displayed. Arrows indicate the location of the gene promoter. (**C**) TF motif enrichment during EpiSC to DE transition. Opened or closed peaks between stages were compared to examine the TF motif enrichments using the one-sided Kolmogorov-Smirnov (KS) test. TF motifs with a KS test effect size ≥ 0.20 (indicated by the dashed lines) are shown. (**D**) Results of *de novo* motif searches within EpiSC or (**E**) MACS-sorted DE ATAC-seq peaks (Homer, v4.11).

Figure 4: Generation of forebrain-patterned organoids from mouse EpiSCs. (**A**) Diagram of the protocol for generation of forebrain-patterned neural organoids from mouse EpiSCs. (**B**) Brightfield images of the organoids at days 1, 5 and 8. Scale bar = 200 μm. (**C**) Confocal immunofluorescence image of developing organoid after 48 hours (d2). Scale bar = 25 μm. (**D-E**) Confocal immunofluorescence images of a d4 organoid. Scale bar = 25 μm. (**F**) Z-stack reconstruction of a d4 organoid. Scale bar = 25 μm. (**G**) Confocal immunofluorescence image of a day 8. Scale bar = 25 μm. For C-G, all organoids were cleared (see materials and methods) and imaged *in toto* without sectioning.

Figure 5: Characterization of the identity of progenitors and neurons generated in forebrain organoids. (**A**) Confocal immunofluorescence image of a d4 organoid. Scale bar = 100 μm. (**B**) Confocal immunofluorescence image of a d8 organoid. Scale bar = 100 μm. (**C**) Diagram of prosomeres in the mouse brain. Modified from (Puelles et al., 2013).

Abbreviations: rostral secondary prosencephalon (RSP), caudal secondary prosencephalon (CSP), prosomeres 1–3 of the diencephalon (p1, p2, and p3), midbrain mesomere 1 and 2 (m1 and m2), isthmic organizer (is), rhombomeres 1–11 (r1 to r11) and spinal cord (SpC). (**D**) Confocal immunofluorescence image of a d8 organoids (further magnified in D′ and D″), immunostained using antibodies against the classical deep layer cortical neuron markers TBR1 and CTIP2, and the forebrain neural progenitor marker OTX2. Scale bars (D) = 100 µm, D′, D″ = 25 µm. (**E**) Confocal immunofluorescence image (3D projection) of a d8 organoid and (**F**) a d12 organoid. Scale bar = 100 µm. (**G**) Confocal immunofluorescence image of a d12 organoid. Scale bar = 100 µm.

Figure 6: Characterization of day 12 forebrain organoids using scRNA-seq. (**A**) UMAP representation of scRNA-sequencing data from 5,652 cells isolated from day 12 organoids. Three broad classes of cell types are highlighted: oligodendrocyte progenitor cells (OPCs), hippocampus/cortical hem, and prethalamus. (**B**) Violin plots of gene expression levels for representative markers used to identify each cluster shown in (**A**) (see also **Supplementary table 5**). (**C**) Annotated UMAP highlighting the expression of *Sox2*, a neuronal progenitor marker, and *Tubb3* (also known as Tuj1), a pan-neuronal marker. (**D**) Cell cycle analyses on d12 scRNA-seq data. (**E**) UMAP representations of cells within putative prethalamic clusters only, highlighting expression of known pre-thalamic markers, from (**A**). (**F**) UMAP representations of gene expression data from only cells within putative hippocampal/cortical hem clusters from (**A**). We observed mutually exclusive expression of *Ctip2* (*Bcl11b*) and *Neurod1*, which delineate the early development of the CA1-2 and CA3 regions of the developing hippocampus, respectively (Simon et al., 2012). (**G**) UMAP representations of gene expression data from only cells within putative OPC clusters from (**A**). (**H**) VoxHunt analyses on the scRNA-seq dataset, separated by cluster, compared to the E18 developing mouse brain. Abbreviations: prosomeres 1–3 of the diencephalon (p1, p2, and p3), midbrain mesomere 1 and 2 (m1 and m2), and rhombomeres 1–11 (r1 to r11).

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1. Systematic optimization of conditions for EpiSC-DE differentiation. (**A**) Overview of state-of-the-art differentiation protocols for human (purple) and mouse (green) PSCs. Further details of each study are provided in **Supplementary Table 1**). (**B-C**) Results of DE protocol optimization experiments. For each condition, immunofluorescence staining was performed following 40 h of differentiation (end of Stage 2). Immunofluorescence images for the indicated antibodies were quantified using CellProfiler. Markers of DE (FOXA2) and mesoderm (CDX2) are shown. (**B**) Summary of results from anterior primitive streak (Stage 1) optimization experiments (n=2 technical replicates). After the indicated conditions for Stage 1, all conditions were exposed to the same Stage 2 conditions, and then quantified. (**C**) Summary of results for optimization of DE commitment (Stage 2) conditions (n=2 technical replicates). For Stage 1, the optimal conditions identified in (**B**) were used, and then cells were switched into the indicated conditions for Stage 2.

Supplementary Figure 2. Extended characterization of EpiSC-derived definitive endoderm. (**A**) Euclidean distance between samples based on ATAC-seq (n=4 per condition, left) and RNA-seq (n=5 for EpiSCs and n=4 for MACs-sorted DE samples, right). (**B**) PCA plot of bulk chromatin accessibility (left) and gene expression (right) of the same C57Bl/6J EpiSC and DE samples as (**A**). (**C**) Unbiased clustered heatmap of the top variable genes as detected by RNA-seq. (**D**) Expression of genes previously identified as DE markers in single cell gene expression studies of hPSC-derived DE and DE isolated from E7.0 embryos (Genga et al., 2019; Nowotschin et al., 2019). (**E**) Representative immunofluorescence images from DE differentiations for a panel of DE markers (FOXA2, GATA6, SOX17, OTX2, EOMES), pluripotency markers (OCT4), and primitive streak/mesoderm markers (T, CDX2). Scale bar = 50 µm.

Supplementary Figure 3: Optimization of conditions for forebrain organoid generation from mouse EpiSCs. (**A**) Brightfield images of d1 organoids in Aggrewell wells. EBs were formed in the presence of Thiazovivin or chroman 1 + emricasan (CE), or the recently published CEPT cocktail (chroman 1, emricasan, polyamine mix, trans-ISRIB) (Chen et al., 2021). Scale bar = 100 µm. (**B**) Confocal immunofluorescence image of a day 4 EB/organoid. Scale bar = 25 µm.

Supplementary Figure 4: EpiSC-derived brain organoids generate prethalamic and cortical-like neuronal populations. (**A,B**) Confocal immunofluorescence images of d8 organoids stained with antibodies against classical

cortical markers (TBR1, TBR2, RELN and PAX6), which are also found in other regions of the brain such as the hippocampus, the cortical hem and the prethalamus, and thalamic markers (TCF7L2). Scale bar = 100 μm. (**C-D**) Confocal immunofluorescence of d8 organoids from C57Bl/6J background. Scale bar = 100 μm. (**E**) Confocal immunofluorescence image of a d12 organoid stained using antibodies against the classical upper-layer cortical marker SATB2. Scale bar = 100 μm. (**F-G**) Confocal immunofluorescence images of d12 organoids. Scale bar = 100 μm.

Supplementary Figure 5: scRNA-seq analyses on day 12 organoids reveals three main cell types. (**A**) Heatmap of the top 10 genes that define each cluster. Abbreviations: VZ = ventricular zone, MZ = mantle zone. (**B**) Expression profile, across clusters, of several prethalamic (*Sp9, Arx, Pax6, Meis2, Ptprd, Islr2*) and axonal (*L1cam, Nrcam*) markers. (**C**) Confocal immunofluorescence image of a day 12 organoid. Scale bar = 30 μm.

Supplementary Table 1: Summary of previous mESC, mEpiLC, and hPSC directed differentiation protocols (related to Figure 2). References used to generate the plot in Fig. 2A (Borowiak et al., 2009; Chen et al., 2013; D'Amour et al., 2005; Diekmann et al., 2019; Green et al., 2011; Kinoshita et al., 2020; Korostylev et al., 2017; Li et al., 2011; Loh et al., 2014; Mfopou et al., 2014; Morrison et al., 2016; Morrison et al., 2008; Mou et al., 2012; Mulas et al., 2017; Ortmann et al., 2020; Sherwood et al., 2011; Shi et al., 2017).

Supplementary Table 2: List of antibodies used.

Supplementary Table 3: List of reagents used.

Supplementary Table 4: Mouse embryonic stem cell lines

Supplementary Table 5: Complete UMAPs, related to figure 6.

Supplementary Video 1. Day 4 organoid stained with antibodies against ZO-1 (green), PH3 (purple) and Nestin (red).

Supplementary Video 2. Live imaging of radial glial-like cells in a d3 organoid. emGFP Sendai virus was added to the EB formation media during the EB formation period to sparsely label developing neuroepithelial progenitor cells. The organoid was then embedded in Matrigel and imaged with a light sheet microscope from day 3 to day 4. The apical membrane is at the bottom of the video and the basal lamina on top.

Supplementary Experimental Methods 1: Step-by-step protocols for directed differentiation experiments

REFERENCES

Arai, Y., and Taverna, E. (2017). Neural Progenitor Cell Polarity and Cortical Development. Front Cell Neurosci 11, 384.

Barry, C., Schmitz, M.T., Jiang, P., Schwartz, M.P., Duffin, B.M., Swanson, S., Bacher, R., Bolin, J.M., Elwell, A.L., McIntosh, B.E., et al. (2017). Species-specific developmental timing is maintained by pluripotent stem cells ex utero. Dev Biol 423, 101-110.

Bernemann, C., Greber, B., Ko, K., Sterneckert, J., Han, D.W., Arauzo-Bravo, M.J., and Scholer, H.R. (2011). Distinct developmental ground states of epiblast stem cell lines determine different pluripotency features. Stem Cells 29, 1496-1503.

Blauwkamp, T.A., Nigam, S., Ardehali, R., Weissman, I.L., and Nusse, R. (2012). Endogenous Wnt signalling in human embryonic stem cells generates an equilibrium of distinct lineage-specified progenitors. Nat Commun 3, 1070.

Borowiak, M., Maehr, R., Chen, S., Chen, A.E., Tang, W., Fox, J.L., Schreiber, S.L., and Melton, D.A. (2009). Small molecules efficiently direct endodermal differentiation of mouse and human embryonic stem cells. Cell Stem Cell 4, 348-358.

Brons, I.G., Smithers, L.E., Trotter, M.W., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R.A., et al. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature 448, 191-195.

Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Curr Protoc Mol Biol 109, 21 29 21-21 29 29.

Caballero, I.M., Manuel, M.N., Molinek, M., Quintana-Urzainqui, I., Mi, D., Shimogori, T., and Price, D.J. (2014). Cell-autonomous repression of Shh by transcription factor Pax6 regulates diencephalic patterning by controlling the central diencephalic organizer. Cell Rep 8, 1405-1418.

Cadwell, C.R., Bhaduri, A., Mostajo-Radji, M.A., Keefe, M.G., and Nowakowski, T.J. (2019). Development and Arealization of the Cerebral Cortex. Neuron 103, 980-1004.

Chen, A.E., Borowiak, M., Sherwood, R.I., Kweudjeu, A., and Melton, D.A. (2013). Functional evaluation of ES cell-derived endodermal populations reveals differences between Nodal and Activin A-guided differentiation. Development 140, 675-686.

Chen, Y., Tristan, C.A., Chen, L., Jovanovic, V.M., Malley, C., Chu, P.H., Ryu, S., Deng, T., Ormanoglu, P., Tao, D., et al. (2021). A versatile polypharmacology platform promotes cytoprotection and viability of human pluripotent and differentiated cells. Nat Methods 18, 528-541.

Churchill, G.A., Gatti, D.M., Munger, S.C., and Svenson, K.L. (2012). The Diversity Outbred mouse population. Mamm Genome 23, 713-718.

Cornacchia, D., Zhang, C., Zimmer, B., Chung, S.Y., Fan, Y., Soliman, M.A., Tchieu, J., Chambers, S.M., Shah, H., Paull, D., et al. (2019). Lipid Deprivation Induces a Stable, Naive-to-Primed Intermediate State of Pluripotency in Human PSCs. Cell Stem Cell 25, 120-136 e110.

Czechanski, A., Byers, C., Greenstein, I., Schrode, N., Donahue, L.R., Hadjantonakis, A.K., and Reinholdt, L.G. (2014). Derivation and characterization of mouse embryonic stem cells from permissive and nonpermissive strains. Nat Protoc 9, 559-574.

D'Amour, K.A., Agulnick, A.D., Eliazer, S., Kelly, O.G., Kroon, E., and Baetge, E.E. (2005). Efficient differentiation of human embryonic stem cells to definitive endoderm. Nat Biotechnol 23, 1534-1541.

Danjo, T., Eiraku, M., Muguruma, K., Watanabe, K., Kawada, M., Yanagawa, Y., Rubenstein, J.L., and Sasai, Y. (2011). Subregional specification of embryonic stem cell-derived ventral telencephalic tissues by timed and combinatory treatment with extrinsic signals. J Neurosci 31, 1919-1933.

Dekkers, J.F., Alieva, M., Wellens, L.M., Ariese, H.C.R., Jamieson, P.R., Vonk, A.M., Amatngalim, G.D., Hu, H., Oost, K.C., Snippert, H.J.G., et al. (2019). High-resolution 3D imaging of fixed and cleared organoids. Nat Protoc 14, 1756-1771.

Diaz-Cuadros, M., Miettinen, T.P., Sheedy, D., Díaz-García, C.M., Gapon, S., Hubaud, A., Yellen, G., Manalis, S.R., Oldham, W., and Pourquié, O. (2021). Metabolic regulation of species-specific developmental rates. bioRxiv, 2021.2008.2027.457974.

Diekmann, U., Wolling, H., Dettmer, R., Niwolik, I., Naujok, O., and Buettner, F.F.R. (2019). Chemically defined and xenogeneic-free differentiation of human pluripotent stem cells into definitive endoderm in 3D culture. Sci Rep 9, 996.

Ebisuya, M., and Briscoe, J. (2018). What does time mean in development? Development 145.

Edri, S., Hayward, P., Baillie-Johnson, P., Steventon, B.J., and Martinez Arias, A. (2019). An epiblast stem cell-derived multipotent progenitor population for axial extension. Development 146.

Eiraku, M., and Sasai, Y. (2011). Mouse embryonic stem cell culture for generation of three-dimensional retinal and cortical tissues. Nat Protoc 7, 69-79.

Eiraku, M., Takata, N., Ishibashi, H., Kawada, M., Sakakura, E., Okuda, S., Sekiguchi, K., Adachi, T., and Sasai, Y. (2011). Self-organizing optic-cup morphogenesis in three-dimensional culture. Nature 472, 51-56.

Eiraku, M., Watanabe, K., Matsuo-Takasaki, M., Kawada, M., Yonemura, S., Matsumura, M., Wataya, T., Nishiyama, A., Muguruma, K., and Sasai, Y. (2008). Self-organized formation of polarized cortical tissues from ESCs and its active manipulation by extrinsic signals. Cell Stem Cell 3, 519-532.

Fleck, J.S., Sanchis-Calleja, F., He, Z., Santel, M., Boyle, M.J., Camp, J.G., and Treutlein, B. (2021). Resolving organoid brain region identities by mapping single-cell genomic data to reference atlases. Cell Stem Cell 28, 1148-1159 e1148.

Fleming, S.J., Marioni, J.C., and Babadi, M. (2019). CellBender remove-background: a deep generative model for unsupervised removal of background noise from scRNA-seq datasets. bioRxiv, 791699.

Fowler, J.L., Ang, L.T., and Loh, K.M. (2020). A critical look: Challenges in differentiating human pluripotent stem cells into desired cell types and organoids. Wiley Interdiscip Rev Dev Biol 9, e368.

Garel, S., Huffman, K.J., and Rubenstein, J.L. (2003). Molecular regionalization of the neocortex is disrupted in Fgf8 hypomorphic mutants. Development 130, 1903-1914.

Gaspar, J.M. (2018). Improved peak-calling with MACS2. bioRxiv, 496521.

Genga, R.M.J., Kernfeld, E.M., Parsi, K.M., Parsons, T.J., Ziller, M.J., and Maehr, R. (2019). Single-Cell RNA-Sequencing-Based CRISPRi Screening Resolves Molecular Drivers of Early Human Endoderm Development. Cell Rep 27, 708-718 e710.

Greber, B., Wu, G., Bernemann, C., Joo, J.Y., Han, D.W., Ko, K., Tapia, N., Sabour, D., Sterneckert, J., Tesar, P., et al. (2010). Conserved and divergent roles of FGF signaling in mouse epiblast stem cells and human embryonic stem cells. Cell Stem Cell 6, 215-226.

Green, M.D., Chen, A., Nostro, M.C., d'Souza, S.L., Schaniel, C., Lemischka, I.R., Gouon-Evans, V., Keller, G., and Snoeck, H.W. (2011). Generation of anterior foregut endoderm from human embryonic and induced pluripotent stem cells. Nat Biotechnol 29, 267-272.

Grove, E.A., Tole, S., Limon, J., Yip, L., and Ragsdale, C.W. (1998). The hem of the embryonic cerebral cortex is defined by the expression of multiple Wnt genes and is compromised in Gli3-deficient mice. Development 125, 2315-2325.

Guo, Q., and Li, J.Y.H. (2019). Defining developmental diversification of diencephalon neurons through single cell gene expression profiling. Development 146.

Hadjantonakis, A.K., Siggia, E.D., and Simunovic, M. (2020). In vitro modeling of early mammalian embryogenesis. Curr Opin Biomed Eng 13, 134-143.

Halliwell, J., Barbaric, I., and Andrews, P.W. (2020). Acquired genetic changes in human pluripotent stem cells: origins and consequences. Nat Rev Mol Cell Biol.

Hanashima, C., Fernandes, M., Hebert, J.M., and Fishell, G. (2007). The role of Foxg1 and dorsal midline signaling in the generation of Cajal-Retzius subtypes. J Neurosci 27, 11103-11111.

Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., 3rd, Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell 184, 3573-3587 e3529.

Hasenpusch-Theil, K., Magnani, D., Amaniti, E.M., Han, L., Armstrong, D., and Theil, T. (2012). Transcriptional analysis of Gli3 mutants identifies Wnt target genes in the developing hippocampus. Cereb Cortex 22, 2878-2893.

Hatami, M., Conrad, S., Naghsh, P., Alvarez-Bolado, G., and Skutella, T. (2018). Cell-Biological Requirements for the Generation of Dentate Gyrus Granule Neurons. Front Cell Neurosci 12, 402.

Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., and Saitou, M. (2011). Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. Cell 146, 519-532.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38, 576-589.

Huch, M., Knoblich, J.A., Lutolf, M.P., and Martinez-Arias, A. (2017). The hope and the hype of organoid research. Development 144, 938-941.

Jouneau, A. (2019). Heterogeneity in Epiblast Stem Cells. Adv Exp Med Biol 1123, 5-17.

Kadoshima, T., Sakaguchi, H., Nakano, T., Soen, M., Ando, S., Eiraku, M., and Sasai, Y. (2013). Self-organization of axial polarity, inside-out layer pattern, and species-specific progenitor dynamics in human ES cell-derived neocortex. Proc Natl Acad Sci U S A 110, 20284-20289.

Kalkan, T., Olova, N., Roode, M., Mulas, C., Lee, H.J., Nett, I., Marks, H., Walker, R., Stunnenberg, H.G., Lilley, K.S., et al. (2017). Tracking the embryonic stem cell transition from ground state pluripotency. Development 144, 1221-1234.

Kelava, I., and Lancaster, M.A. (2016). Stem Cell Models of Human Brain Development. Cell Stem Cell 18, 736-748.

Kessaris, N., Fogarty, M., Iannarelli, P., Grist, M., Wegner, M., and Richardson, W.D. (2006). Competing waves of oligodendrocytes in the forebrain and postnatal elimination of an embryonic lineage. Nat Neurosci 9, 173-179.

Kim, D.W., Washington, P.W., Wang, Z.Q., Lin, S.H., Sun, C., Ismail, B.T., Wang, H., Jiang, L., and Blackshaw, S. (2020). The cellular and molecular landscape of hypothalamic patterning and differentiation from embryonic to late postnatal development. Nat Commun 11, 4360.

Kinoshita, M., Barber, M., Mansfield, W., Cui, Y., Spindlow, D., Stirparo, G.G., Dietmann, S., Nichols, J., and Smith, A. (2020). Capture of mouse and human stem cells with features of formative pluripotency. bioRxiv, 2020.2009.2004.283218.

Kleiman, R.J., and Engle, S.J. (2021). Human inducible pluripotent stem cells: Realization of initial promise in drug discovery. Cell Stem Cell 28, 1507-1515.

Kojima, Y., Kaufman-Francis, K., Studdert, J.B., Steiner, K.A., Power, M.D., Loebel, D.A., Jones, V., Hor, A., de Alencastro, G., Logan, G.J., et al. (2014). The transcriptional and functional properties of mouse epiblast stem cells resemble the anterior primitive streak. Cell Stem Cell 14, 107-120.

Korostylev, A., Mahaddalkar, P.U., Keminer, O., Hadian, K., Schorpp, K., Gribbon, P., and Lickert, H. (2017). A high-content small molecule screen identifies novel inducers of definitive endoderm. Mol Metab 6, 640-650.

Kurek, D., Neagu, A., Tastemel, M., Tuysuz, N., Lehmann, J., van de Werken, H.J.G., Philipsen, S., van der Linden, R., Maas, A., van, I.W.F.J., et al. (2015). Endogenous WNT signals mediate BMP-induced and spontaneous differentiation of epiblast stem cells and human embryonic stem cells. Stem Cell Reports 4, 114-128.

Lager, A.M., Corradin, O.G., Cregg, J.M., Elitt, M.S., Shick, H.E., Clayton, B.L.L., Allan, K.C., Olsen, H.E., Madhavan, M., and Tesar, P.J. (2018). Rapid functional genetics of the oligodendrocyte lineage using pluripotent stem cells. Nat Commun 9, 3708.

Lancaster, M.A., Renner, M., Martin, C.A., Wenzel, D., Bicknell, L.S., Hurles, M.E., Homfray, T., Penninger, J.M., Jackson, A.P., and Knoblich, J.A. (2013). Cerebral organoids model human brain development and microcephaly. Nature 501, 373-379.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357-359.

Li, F., He, Z., Li, Y., Liu, P., Chen, F., Wang, M., Zhu, H., Ding, X., Wangensteen, K.J., Hu, Y., et al. (2011). Combined activin A/LiCl/Noggin treatment improves production of mouse embryonic stem cell-derived definitive endoderm cells. J Cell Biochem 112, 1022-1034.

Li, L.C., Qiu, W.L., Zhang, Y.W., Xu, Z.R., Xiao, Y.N., Hou, C., Lamaoqiezhong, Yu, P., Cheng, X., and Xu, C.R. (2018). Single-cell transcriptomic analyses reveal distinct dorsal/ventral pancreatic programs. EMBO Rep 19.

Li, Q.V., Dixon, G., Verma, N., Rosen, B.P., Gordillo, M., Luo, R., Xu, C., Wang, Q., Soh, C.L., Yang, D., et al. (2019). Genome-scale screens identify JNK-JUN signaling as a barrier for pluripotency exit and endoderm differentiation. Nat Genet 51, 999-1010.

Li, Y., Lopez-Huerta, V.G., Adiconis, X., Levandowski, K., Choi, S., Simmons, S.K., Arias-Garcia, M.A., Guo, B., Yao, A.Y., Blosser, T.R., et al. (2020). Distinct subnetworks of the thalamic reticular nucleus. Nature 583, 819-824.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923-930.

Loh, K.M., Ang, L.T., Zhang, J., Kumar, V., Ang, J., Auyeong, J.Q., Lee, K.L., Choo, S.H., Lim, C.Y., Nichane, M., et al. (2014). Efficient endoderm induction from human pluripotent stem cells by logically directing signals controlling lineage bifurcations. Cell Stem Cell 14, 237-252.

Loh, K.M., Chen, A., Koh, P.W., Deng, T.Z., Sinha, R., Tsai, J.M., Barkal, A.A., Shen, K.Y., Jain, R., Morganti, R.M., et al. (2016). Mapping the Pairwise Choices Leading from Pluripotency to Human Bone, Heart, and Other Mesoderm Cell Types. Cell 166, 451-467.

Loh, K.M., Lim, B., and Ang, L.T. (2015). Ex uno plures: molecular designs for embryonic pluripotency. Physiol Rev 95, 245-295.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 550.

Lusis, A.J., Seldin, M.M., Allayee, H., Bennett, B.J., Civelek, M., Davis, R.C., Eskin, E., Farber, C.R., Hui, S., Mehrabian, M., et al. (2016). The Hybrid Mouse Diversity Panel: a resource for systems genetics analyses of metabolic and cardiovascular traits. J Lipid Res 57, 925-942.

Mandai, K., Reimert, D.V., and Ginty, D.D. (2014). Linx mediates interaxonal interactions and formation of the internal capsule. Neuron 83, 93-103.

Manuel, M.N., Mi, D., Mason, J.O., and Price, D.J. (2015). Regulation of cerebral cortical neurogenesis by the Pax6 transcription factor. Front Cell Neurosci 9, 70.

Marini, F., and Binder, H. (2019). pcaExplorer: an R/Bioconductor package for interacting with RNA-seq principal components. BMC Bioinformatics 20, 331.

Martinez-Ferre, A., and Martinez, S. (2009). The development of the thalamic motor learning area is regulated by Fgf8 expression. J Neurosci 29, 13389-13400.

Matsuda, M., Hayashi, H., Garcia-Ojalvo, J., Yoshioka-Kobayashi, K., Kageyama, R., Yamanaka, Y., Ikeya, M., Toguchida, J., Alev, C., and Ebisuya, M. (2020). Species-specific segmentation clock periods are due to differential biochemical reaction speeds. Science 369, 1450-1455.

Mezger, A., Klemm, S., Mann, I., Brower, K., Mir, A., Bostick, M., Farmer, A., Fordyce, P., Linnarsson, S., and Greenleaf, W. (2018). High-throughput chromatin accessibility profiling at single-cell resolution. Nat Commun 9, 3647.

Mfopou, J.K., Geeraerts, M., Dejene, R., Van Langenhoven, S., Aberkane, A., Van Grunsven, L.A., and Bouwens, L. (2014). Efficient definitive endoderm induction from mouse embryonic stem cell adherent cultures: a rapid screening model for differentiation studies. Stem Cell Res 12, 166-177.

Molnar, Z., Garel, S., Lopez-Bendito, G., Maness, P., and Price, D.J. (2012). Mechanisms controlling the guidance of thalamocortical axons through the embryonic forebrain. Eur J Neurosci 35, 1573-1585.

Montiel, J.F., and Aboitiz, F. (2015). Pallial patterning and the origin of the isocortex. Front Neurosci 9, 377.

Morgani, S., Nichols, J., and Hadjantonakis, A.K. (2017). The many faces of Pluripotency: in vitro adaptations of a continuum of in vivo states. BMC Dev Biol 17, 7.

Morgani, S.M., and Hadjantonakis, A.K. (2020). Signaling regulation during gastrulation: Insights from mouse embryos and in vitro systems. Curr Top Dev Biol 137, 391-431.

Morgani, S.M., Metzger, J.J., Nichols, J., Siggia, E.D., and Hadjantonakis, A.K. (2018). Micropattern differentiation of mouse pluripotent stem cells recapitulates embryo regionalized cell fate patterning. Elife 7.

Moris, N., Martinez Arias, A., and Steventon, B. (2020). Experimental embryology of gastrulation: pluripotent stem cells as a new model system. Curr Opin Genet Dev 64, 78-83.

Morrison, G., Scognamiglio, R., Trumpp, A., and Smith, A. (2016). Convergence of cMyc and beta-catenin on Tcf7l1 enables endoderm specification. EMBO J 35, 356-368.

Morrison, G.M., Oikonomopoulou, I., Migueles, R.P., Soneji, S., Livigni, A., Enver, T., and Brickman, J.M. (2008). Anterior definitive endoderm from ESCs reveals a role for FGF signaling. Cell Stem Cell 3, 402-415.

Mou, H., Zhao, R., Sherwood, R., Ahfeldt, T., Lapey, A., Wain, J., Sicilian, L., Izvolsky, K., Musunuru, K., Cowan, C., et al. (2012). Generation of multipotent lung and airway progenitors from mouse ESCs and patient-specific cystic fibrosis iPSCs. Cell Stem Cell 10, 385-397.

Mulas, C., Kalkan, T., and Smith, A. (2017). NODAL Secures Pluripotency upon Embryonic Stem Cell Progression from the Ground State. Stem Cell Reports 9, 77-91.

Nagalski, A., Puelles, L., Dabrowski, M., Wegierski, T., Kuznicki, J., and Wisniewska, M.B. (2016). Molecular anatomy of the thalamic complex and the underlying transcription factors. Brain Struct Funct 221, 2493-2510.

Najm, F.J., Zaremba, A., Caprariello, A.V., Nayak, S., Freundt, E.C., Scacheri, P.C., Miller, R.H., and Tesar, P.J. (2011). Rapid and robust generation of functional oligodendrocyte progenitor cells from epiblast stem cells. Nat Methods 8, 957-962.

Nasu, M., Takata, N., Danjo, T., Sakaguchi, H., Kadoshima, T., Futaki, S., Sekiguchi, K., Eiraku, M., and Sasai, Y. (2012). Robust formation and maintenance of continuous stratified cortical neuroepithelium by laminin-containing matrix in mouse ES cell culture. PLoS One 7, e53024.

Nichols, J., and Smith, A. (2012). Pluripotency in the embryo and in culture. Cold Spring Harb Perspect Biol 4, a008128.

Noctor, S.C., Martinez-Cerdeno, V., Ivic, L., and Kriegstein, A.R. (2004). Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. Nat Neurosci 7, 136-144.

Nowotschin, S., Setty, M., Kuo, Y.Y., Liu, V., Garg, V., Sharma, R., Simon, C.S., Saiz, N., Gardner, R., Boutet, S.C., et al. (2019). The emergent landscape of the mouse gut endoderm at single-cell resolution. Nature 569, 361-367.

Ono, K., Clavairoly, A., Nomura, T., Gotoh, H., Uno, A., Armant, O., Takebayashi, H., Zhang, Q., Shimamura, K., Itohara, S., et al. (2014). Development of the prethalamus is crucial for thalamocortical projection formation and is regulated by Olig2. Development 141, 2075-2084.

Ortmann, D., Brown, S., Czechanski, A., Aydin, S., Muraro, D., Huang, Y., Tomaz, R.A., Osnato, A., Canu, G., Wesley, B.T., et al. (2020). Naive Pluripotent Stem Cells Exhibit Phenotypic Variability that Is Driven by Genetic Variation. Cell Stem Cell 27, 470-481 e476.

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 14, 417-419.

Puelles, L., Harrison, M., Paxinos, G., and Watson, C. (2013). A developmental ontology for the mammalian brain based on the prosomeric model. Trends Neurosci 36, 570-578.

Puelles, L., Kuwana, E., Puelles, E., Bulfone, A., Shimamura, K., Keleher, J., Smiga, S., and Rubenstein, J.L. (2000). Pallial and subpallial derivatives in the embryonic chick and mouse telencephalon, traced by the expression of the genes Dlx-2, Emx-1, Nkx-2.1, Pax-6, and Tbr-1. J Comp Neurol 424, 409-438.

Qian, X., Davis, A.A., Goderie, S.K., and Temple, S. (1997). FGF2 concentration regulates the generation of neurons and glia from multipotent cortical stem cells. Neuron 18, 81-93.

Qian, X., Jacob, F., Song, M.M., Nguyen, H.N., Song, H., and Ming, G.L. (2018). Generation of human brain region-specific organoids using a miniaturized spinning bioreactor. Nat Protoc 13, 565-580.

Quintana-Urzainqui, I., Hernandez-Malmierca, P., Clegg, J.M., Li, Z., Kozic, Z., and Price, D.J. (2020). The role of the diencephalon in the guidance of thalamocortical axons in mice. Development 147.

Rayon, T., Stamataki, D., Perez-Carrasco, R., Garcia-Perez, L., Barrington, C., Melchionda, M., Exelby, K., Lazaro, J., Tybulewicz, V.L.J., Fisher, E.M.C., et al. (2020). Species-specific pace of development is associated with differences in protein stability. Science 369.

Renner, M., Lancaster, M.A., Bian, S., Choi, H., Ku, T., Peer, A., Chung, K., and Knoblich, J.A. (2017). Self-organized developmental patterning and differentiation in cerebral organoids. EMBO J 36, 1316-1329.

Rifes, P., Isaksson, M., Rathore, G.S., Aldrin-Kirk, P., Moller, O.K., Barzaghi, G., Lee, J., Egerod, K.L., Rausch, D.M., Parmar, M., et al. (2020). Modeling neural tube development by differentiation of human embryonic stem cells in a microfluidic WNT gradient. Nat Biotechnol 38, 1265-1273.

Rossant, J. (2015). Mouse and human blastocyst-derived stem cells: vive les differences. Development 142, 9-12.

Rossant, J., and Tam, P.P.L. (2017). New Insights into Early Human Development: Lessons for Stem Cell Derivation and Differentiation. Cell Stem Cell 20, 18-28.

Schlissel, G., and Li, P. (2020). Synthetic Developmental Biology: Understanding Through Reconstitution. Annu Rev Cell Dev Biol 36, 339-357.

Sestan, N., and State, M.W. (2018). Lost in Translation: Traversing the Complex Path from Genomics to Therapeutics in Autism Spectrum Disorder. Neuron 100, 406-423.

Sherwood, R.I., Maehr, R., Mazzoni, E.O., and Melton, D.A. (2011). Wnt signaling specifies and patterns intestinal endoderm. Mech Dev 128, 387-400.

Shi, Z.D., Lee, K., Yang, D., Amin, S., Verma, N., Li, Q.V., Zhu, Z., Soh, C.L., Kumar, R., Evans, T., et al. (2017). Genome Editing in hPSCs Reveals GATA6 Haploinsufficiency and a Genetic Interaction with GATA4 in Human Pancreatic Development. Cell Stem Cell 20, 675-688 e676.

Shimamura, K., and Rubenstein, J.L. (1997). Inductive interactions direct early regionalization of the mouse forebrain. Development 124, 2709-2718.

Shiraishi, A., Muguruma, K., and Sasai, Y. (2017). Generation of thalamic neurons from mouse embryonic stem cells. Development 144, 1211-1220.

Simon, R., Brylka, H., Schwegler, H., Venkataramanappa, S., Andratschke, J., Wiegreffe, C., Liu, P., Fuchs, E., Jenkins, N.A., Copeland, N.G., et al. (2012). A dual function of Bcl11b/Ctip2 in hippocampal neurogenesis. EMBO J 31, 2922-2936.

Skelly, D.A., Czechanski, A., Byers, C., Aydin, S., Spruce, C., Olivier, C., Choi, K., Gatti, D.M., Raghupathy, N., Keele, G.R., et al. (2020). Mapping the Effects of Genetic Variation on Chromatin State and Gene Expression Reveals Loci That Control Ground State Pluripotency. Cell Stem Cell 27, 459-469 e458.

Smith, A. (2017). Formative pluripotency: the executive phase in a developmental continuum. Development 144, 365-373.

Sommer, L., Rao, M., and Anderson, D.J. (1997). RPTP delta and the novel protein tyrosine phosphatase RPTP psi are expressed in restricted regions of the developing central nervous system. Dev Dyn 208, 48-61.

Song, L., Chen, J., Peng, G., Tang, K., and Jing, N. (2016). Dynamic Heterogeneity of Brachyury in Mouse Epiblast Stem Cells Mediates Distinct Response to Extrinsic Bone Morphogenetic Protein (BMP) Signaling. J Biol Chem 291, 15212-15225.

Storm, E.E., Garel, S., Borello, U., Hebert, J.M., Martinez, S., McConnell, S.K., Martin, G.R., and Rubenstein, J.L. (2006). Dose-dependent functions of Fgf8 in regulating telencephalic patterning centers. Development 133, 1831-1844.

Strawbridge, S.E., Blanchard, G.B., Smith, A., Kugler, H., and Martello, G. (2020). Embryonic stem cells commit to differentiation by symmetric divisions following a variable lag period. bioRxiv, 2020.2006.2017.157578.

Sumi, T., Oki, S., Kitajima, K., and Meno, C. (2013). Epiblast ground state is controlled by canonical Wnt/beta-catenin signaling in the postimplantation mouse embryo and epiblast stem cells. PLoS One 8, e63378.

Takata, N., Sakakura, E., Eiraku, M., Kasukawa, T., and Sasai, Y. (2017). Self-patterning of rostral-caudal neuroectoderm requires dual role of Fgf signaling for localized Wnt antagonism. Nat Commun 8, 1339.

Takiguchi-Hayashi, K., Sekiguchi, M., Ashigaki, S., Takamatsu, M., Hasegawa, H., Suzuki-Migishima, R., Yokoyama, M., Nakanishi, S., and Tanabe, Y. (2004). Generation of reelin-positive marginal zone cells from the caudomedial wall of telencephalic vesicles. J Neurosci 24, 2286-2295.

Tchieu, J., Zimmer, B., Fattahi, F., Amin, S., Zeltner, N., Chen, S., and Studer, L. (2017). A Modular Platform for Differentiation of Human PSCs into All Major Ectodermal Lineages. Cell Stem Cell 21, 399-410 e397.

Tekki-Kessaris, N., Woodruff, R., Hall, A.C., Gaffield, W., Kimura, S., Stiles, C.D., Rowitch, D.H., and Richardson, W.D. (2001). Hedgehog-dependent oligodendrocyte lineage specification in the telencephalon. Development 128, 2545-2554.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature 448, 196-199.

Thompson, C.L., Ng, L., Menon, V., Martinez, S., Lee, C.K., Glattfelder, K., Sunkin, S.M., Henry, A., Lau, C., Dang, C., et al. (2014). A high-resolution spatiotemporal atlas of gene expression of the developing mouse brain. Neuron 83, 309-323.

Tsakiridis, A., Huang, Y., Blin, G., Skylaki, S., Wymeersch, F., Osorno, R., Economou, C., Karagianni, E., Zhao, S., Lowell, S., et al. (2014). Distinct Wnt-driven primitive streak-like populations reflect in vivo lineage precursors. Development 141, 1209-1221.

Tuttle, R., Nakagawa, Y., Johnson, J.E., and O'Leary, D.D. (1999). Defects in thalamocortical axon pathfinding correlate with altered cell domains in Mash-1-deficient mice. Development 126, 1903-1916.

Vallier, L., Touboul, T., Chng, Z., Brimpari, M., Hannan, N., Millan, E., Smithers, L.E., Trotter, M., Rugg-Gunn, P., Weber, A., et al. (2009). Early cell fate decisions of human embryonic stem cells and mouse epiblast stem cells are controlled by the same signalling pathways. PLoS One 4, e6082.

Vieira, C., Pombero, A., Garcia-Lopez, R., Gimeno, L., Echevarria, D., and Martinez, S. (2010). Molecular mechanisms controlling brain development: an overview of neuroepithelial secondary organizers. Int J Dev Biol 54, 7-20.

Virolainen, S.M., Achim, K., Peltopuro, P., Salminen, M., and Partanen, J. (2012). Transcriptional regulatory mechanisms underlying the GABAergic neuron fate in different diencephalic prosomeres. Development 139, 3795-3805.

Voelkl, B., Altman, N.S., Forsman, A., Forstmeier, W., Gurevitch, J., Jaric, I., Karp, N.A., Kas, M.J., Schielzeth, H., Van de Casteele, T., et al. (2020). Reproducibility of animal research in light of biological variation. Nat Rev Neurosci 21, 384-393.

Watanabe, M., Haney, J.R., Vishlaghi, N., Turcios, F., Buth, J.E., Gu, W., Collier, A.J., Miranda, O.A., Chen, D., Sabri, S., et al. (2019). TGFβ superfamily signaling regulates the state of human stem cell pluripotency and competency to create telencephalic organoids. bioRxiv, 2019.2012.2013.875773.

Wataya, T., Ando, S., Muguruma, K., Ikeda, H., Watanabe, K., Eiraku, M., Kawada, M., Takahashi, J., Hashimoto, N., and Sasai, Y. (2008). Minimization of exogenous signals in ES cell culture induces rostral hypothalamic differentiation. Proc Natl Acad Sci U S A 105, 11796-11801.

Weinberger, L., Ayyash, M., Novershtern, N., and Hanna, J.H. (2016). Dynamic stem cell states: naive to primed pluripotency in rodents and humans. Nat Rev Mol Cell Biol 17, 155-169.

Wu, J., Okamura, D., Li, M., Suzuki, K., Luo, C., Ma, L., He, Y., Li, Z., Benner, C., Tamura, I., et al. (2015). An alternative pluripotent state confers interspecies chimaeric competency. Nature 521, 316-321.

Xiang, Y., Tanaka, Y., Cakir, B., Patterson, B., Kim, K.Y., Sun, P., Kang, Y.J., Zhong, M., Liu, X., Patra, P., et al. (2019). hESC-Derived Thalamic Organoids Form Reciprocal Projections When Fused with Cortical Organoids. Cell Stem Cell 24, 487-497 e487.

Xiang, Y., Zhang, Y., Xu, Q., Zhou, C., Liu, B., Du, Z., Zhang, K., Zhang, B., Wang, X., Gayen, S., et al. (2020). Epigenomic analysis of gastrulation identifies a unique chromatin state for primed pluripotency. Nat Genet 52, 95-105.

Yiangou, L., Ross, A.D.B., Goh, K.J., and Vallier, L. (2018). Human Pluripotent Stem Cell-Derived Endoderm for Modeling Development and Clinical Applications. Cell Stem Cell 22, 485-499.

Yu, L., Wei, Y., Sun, H.X., Mahdi, A.K., Pinzon Arteaga, C.A., Sakurai, M., Schmitz, D.A., Zheng, C., Ballard, E.D., Li, J., et al. (2021). Derivation of Intermediate Pluripotent Stem Cells Amenable to Primordial Germ Cell Specification. Cell Stem Cell 28, 550-567 e512.

Zorn, A.M., and Wells, J.M. (2009). Vertebrate endoderm development and organ formation. Annu Rev Cell Dev Biol 25, 221-251.

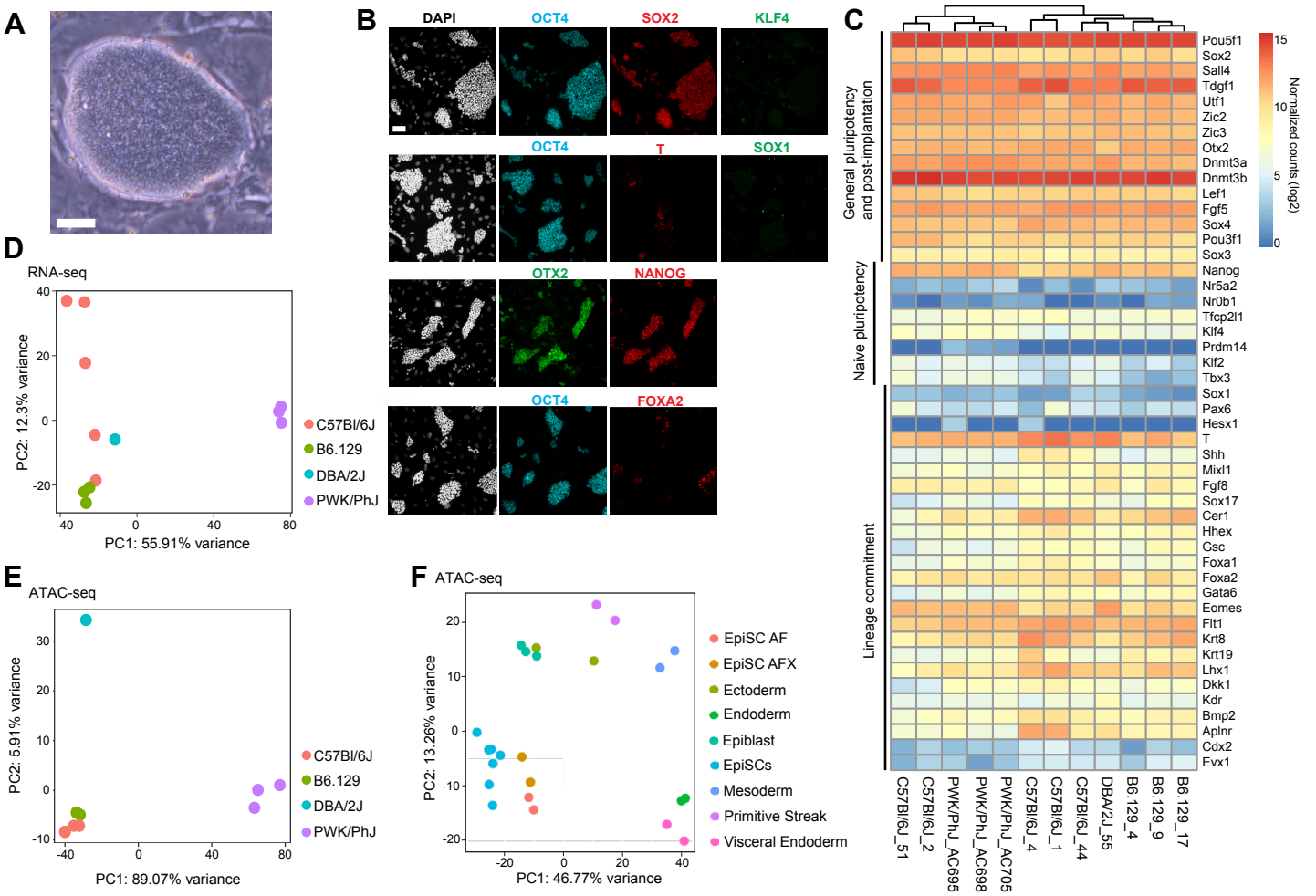


Figure 1: Derivation and characterization of mouse EpiSCs cultured in primed ground state conditions. (A) Phase contrast image of an EpiSC colony. Scale bar = 50µm. **(B)** Immunofluorescence images of EpiSC cultures. Markers of general pluripotency (SOX2, OCT4), primed pluripotency (OTX2), naive pluripotency (KFL4, NANOG), and lineage commitment (T, FOXA2, SOX1). Scale bar = 50 µm. **(C)** Gene expression heatmap of RNA-sequencing data from EpiSC lines (n=12) for selected genes associated with different stages of pluripotency and lineage-commitment. Each EpiSC line was derived from a distinct mouse ESC line via in vitro naïve-to-primed conversion, and represent biological replicates. **(D)** Principal component analysis (PCA) plot of RNA-seq data (n =12 biological replicates; C57Bl/6J = 5, B6.129 = 3, PWK/PhJ = 3, DBA/2J = 1) and **(E)** ATAC-seq data (n =9 biological replicates; C57Bl/6J = 3, B6.129 = 2, PWK/PhJ = 3, DBA/2J = 1) for EpiSC lines. **(F)** PCA plot of ATAC-seq data for EpiSC lines (n =7 biological replicates; C57Bl/6J = 4, PWK/PhJ = 2, DBA/2J = 1). For comparison, ATAC-seq data from an additional set of EpiSC lines EpiSC-AF (cultured with Activin A and Fgf2) and EpiSC-AFX (cultured with Activin A, Fgf2, and Wnt inhibitor XAV939) from (Kinoshita et al., 2020) and ATAC-seq data from several early embryonic lineages *in vivo* (E6.5 epiblast, E6.5 visceral endoderm, and E7.5 ectoderm, E7.5 mesoderm, E7.5 endoderm, and E7.5 primitive streak were included from (Xiang et al., 2020).

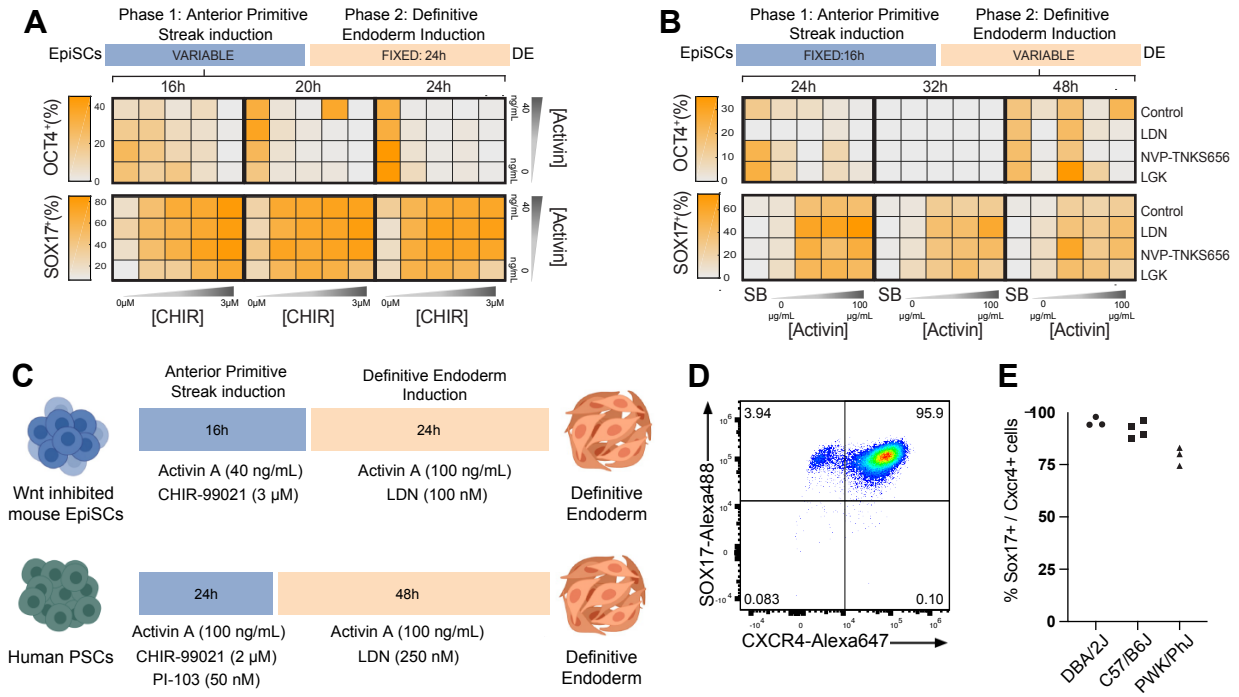


Figure 2: Systematic optimization of conditions for EpiSC-DE differentiation. (A-B) Results of DE protocol optimization experiments. For each condition, immunofluorescence staining was performed following 40 h of differentiation (end of Stage 2). Immunofluorescence images for the indicated antibodies were quantified using CellProfiler. Markers of DE (FOXA2, SOX17), pluripotency (OCT4), and PS/mesoderm (T, CDX2) were examined. **(A)** Summary of results from anterior primitive streak (Stage 1) optimization experiments (n=2 technical replicates). **(B)** Summary of results for optimization of DE commitment (Stage 2) conditions (n=2 technical replicates). For Stage 1, the optimal conditions identified in **(A)** were used, and then cells were switched into the indicated conditions for Stage 2. **(C)** Diagram of the optimized protocol for differentiation of mouse EpiSCs into DE and a current human PSC to DE protocol (from Loh et al, 2014). **(D)** FACS Quantification of immunostaining for DE markers (SOX17-Alexa488 and CXCR4-Alexa647) from cells at the endpoint of the optimized DE differentiation protocol. **(E)** Quantification of DE differentiation purity (SOX17/CXCR4 double positive cells) across multiple experiments performed using the new optimized DE differentiation protocol. Each dot indicates a technical replicate from the indicated line (C57Bl/6J, n=4; DBA/2J, n=3; PWK/PhJ, n=3).

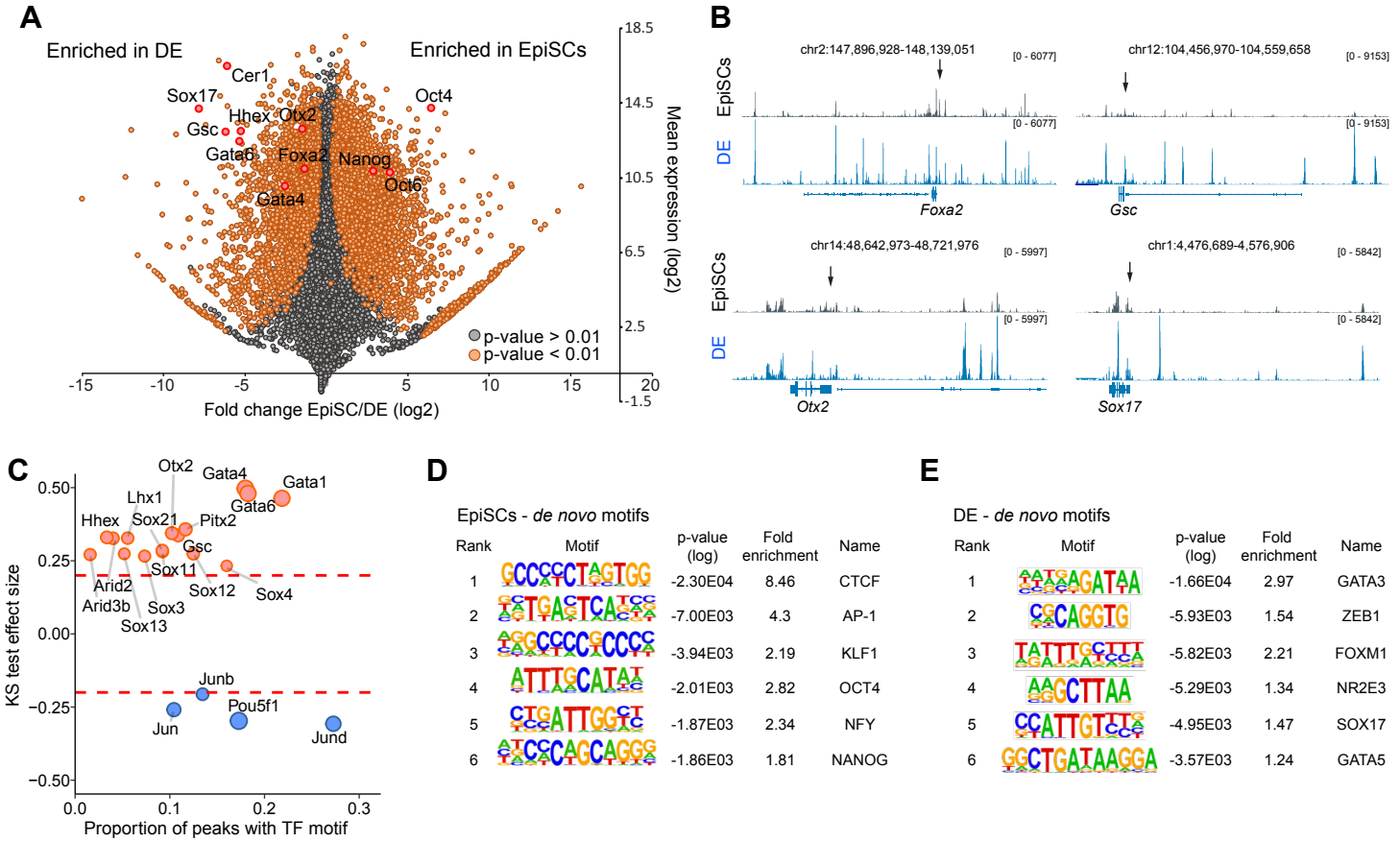


Figure 3: Genomic characterization of EpiSC-derived definitive endoderm. (A) RNA-sequencing data comparing gene expression between C57Bl/6J EpiSCs (n=5, biological replicates) and C57Bl/6J MACS-sorted DE (n=4, biological replicates). Genes exhibiting significant differences in expression between EpiSCs and DE are highlighted in orange ($P < 0.01$, Wald test and corrected using the Benjamini-Hochberg method). A set of critical genes for EpiSCs and DE are further annotated in red. **(B)** Chromatin accessibility (ATAC-seq) data from EpiSC (top row) or DE (bottom row) samples. Loci of genes known to play critical roles in DE development are displayed. Arrows indicate the location of the gene promoter. **(C)** TF motif enrichment during EpiSC to DE transition. Opened or closed peaks between stages were compared to examine the TF motif enrichments using the one-sided Kolmogorov-Smirnov (KS) test. TF motifs with a KS test effect size ≥ 0.20 (indicated by the dashed lines) are shown. **(D)** Results of *de novo* motif searches within EpiSC or **(E)** MACS-sorted DE ATAC-seq peaks (Homer, v4.11).

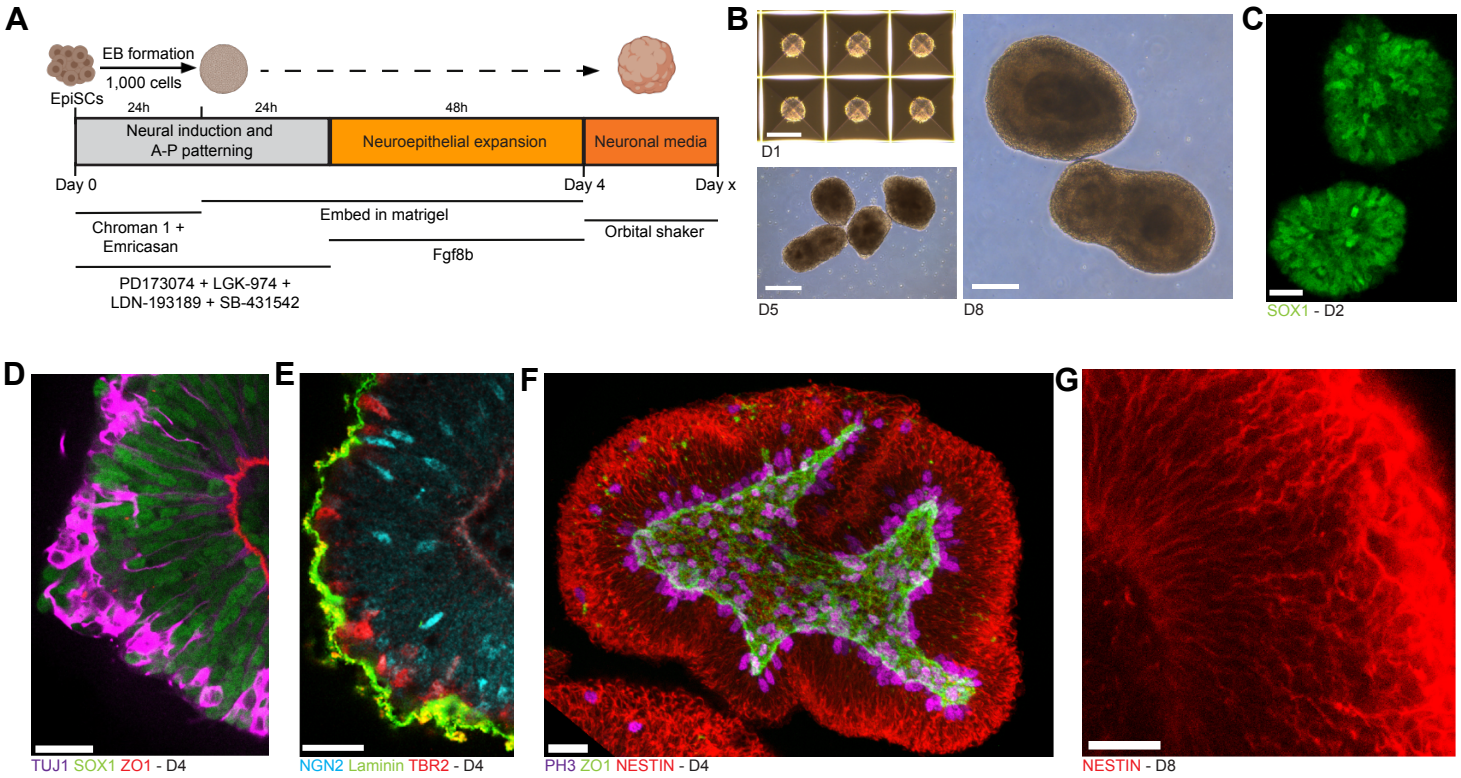


Figure 4: Generation of forebrain-patterned organoids from mouse EpiSCs. (A) Diagram of the protocol for generation of forebrain-patterned neural organoids from mouse EpiSCs. **(B)** Brightfield images of the organoids at days 1, 5 and 8. Scale bar = 200 µm. **(C)** Confocal immunofluorescence image of developing organoid after 48 hours (d2). Scale bar = 25 µm. **(D-E)** Confocal immunofluorescence images of a d4 organoid. Scale bar = 25 µm. **(F)** Z-stack reconstruction of a d4 organoid. Scale bar = 25 µm. **(G)** Confocal immunofluorescence image (z-slice) of a day 8 organoid. Scale bar = 25 µm. For C-G, all organoids were cleared (see materials and methods) and imaged *in toto* without sectioning.

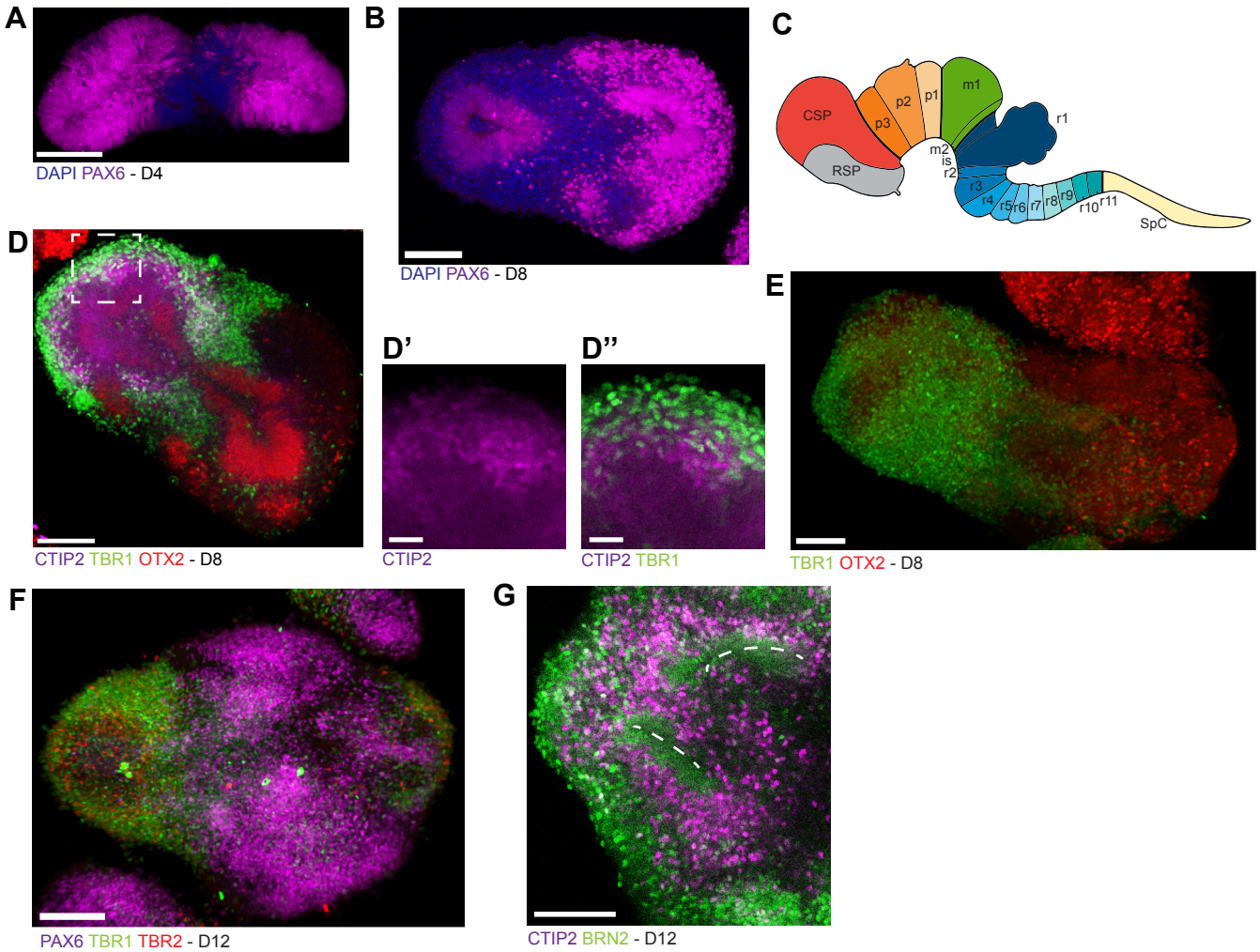


Figure 5: Characterization of the identity of progenitors and neurons generated in forebrain organoids. (A) Confocal immunofluorescence image of a d4 organoid. Scale bar = 100 µm. **(B)** Confocal immunofluorescence image of a d8 organoid. Scale bar = 100 µm. **(C)** Diagram of prosomeres in the mouse brain. Modified from (Puelles et al., 2013). Abbreviations: rostral secondary prosencephalon (RSP), caudal secondary prosencephalon (CSP), prosomeres 1–3 of the diencephalon (p1, p2, and p3), midbrain mesomere 1 and 2 (m1 and m2), isthmic organizer (is), rhombomeres 1–11 (r1 to r11) and spinal cord (SpC). **(D)** Confocal immunofluorescence image of a d8 organoid (further magnified in **D'** and **D"**), stained using antibodies against the classical deep layer cortical markers TBR1 and CTIP2, and OTX2 to stain the neuroepithelia. Scale bars **(D)** = 100 µm, **D', D"** = 25 µm. **(E)** Confocal immunofluorescence image (3D projection) of a d8 organoid and **(F)** a d12 organoid. Scale bar = 100 µm. **(G)** Confocal immunofluorescence image of a d12 organoid. Scale bar = 100 µm.

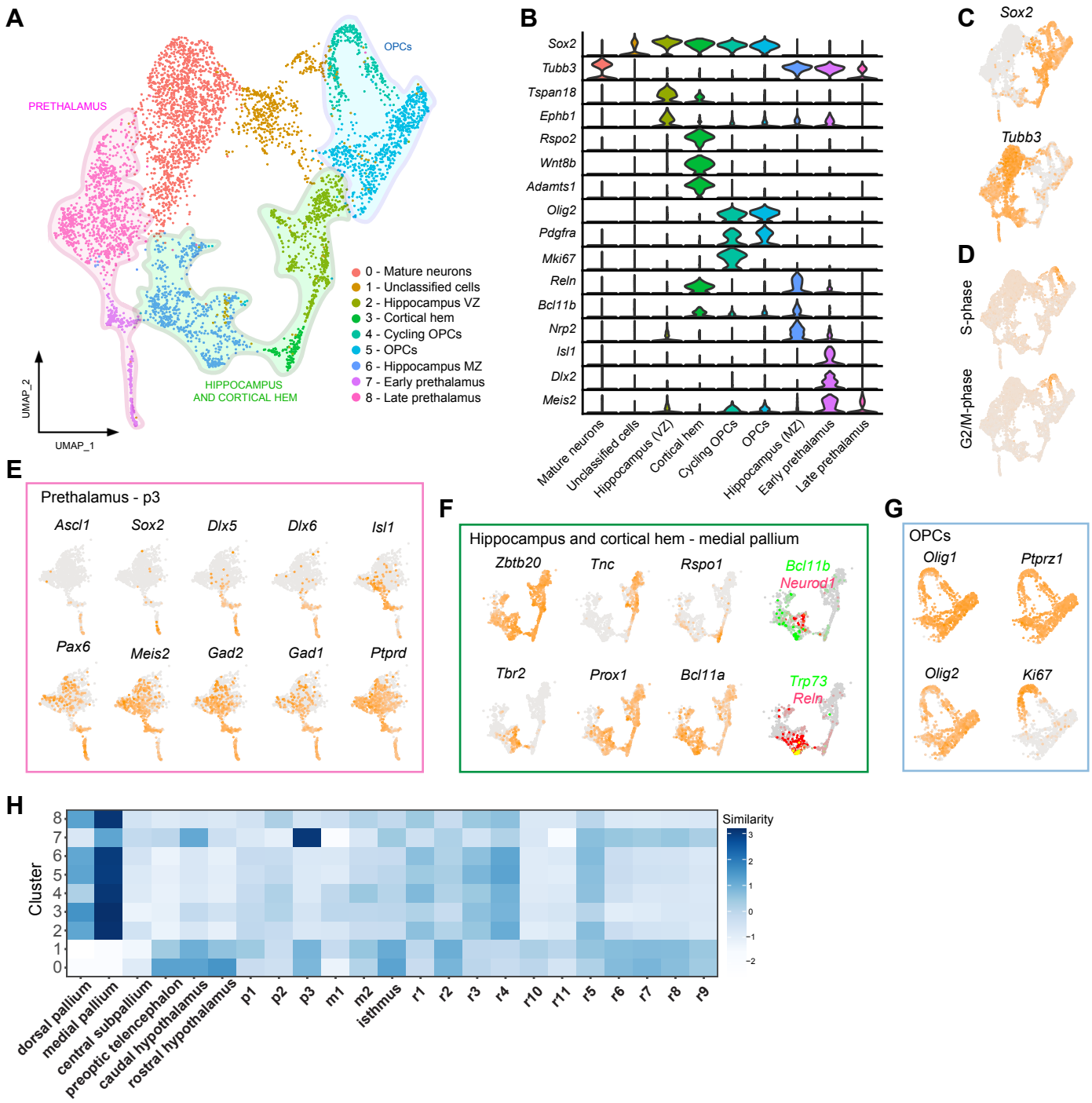
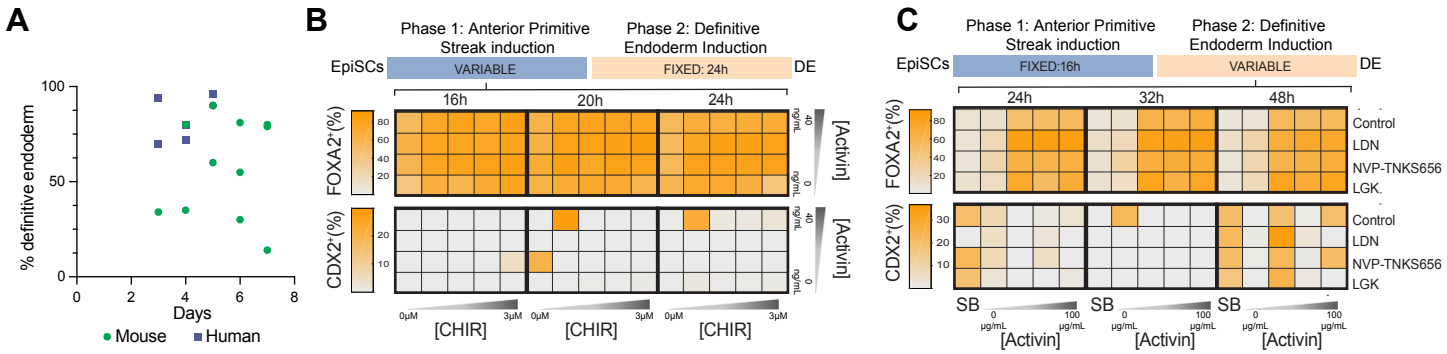
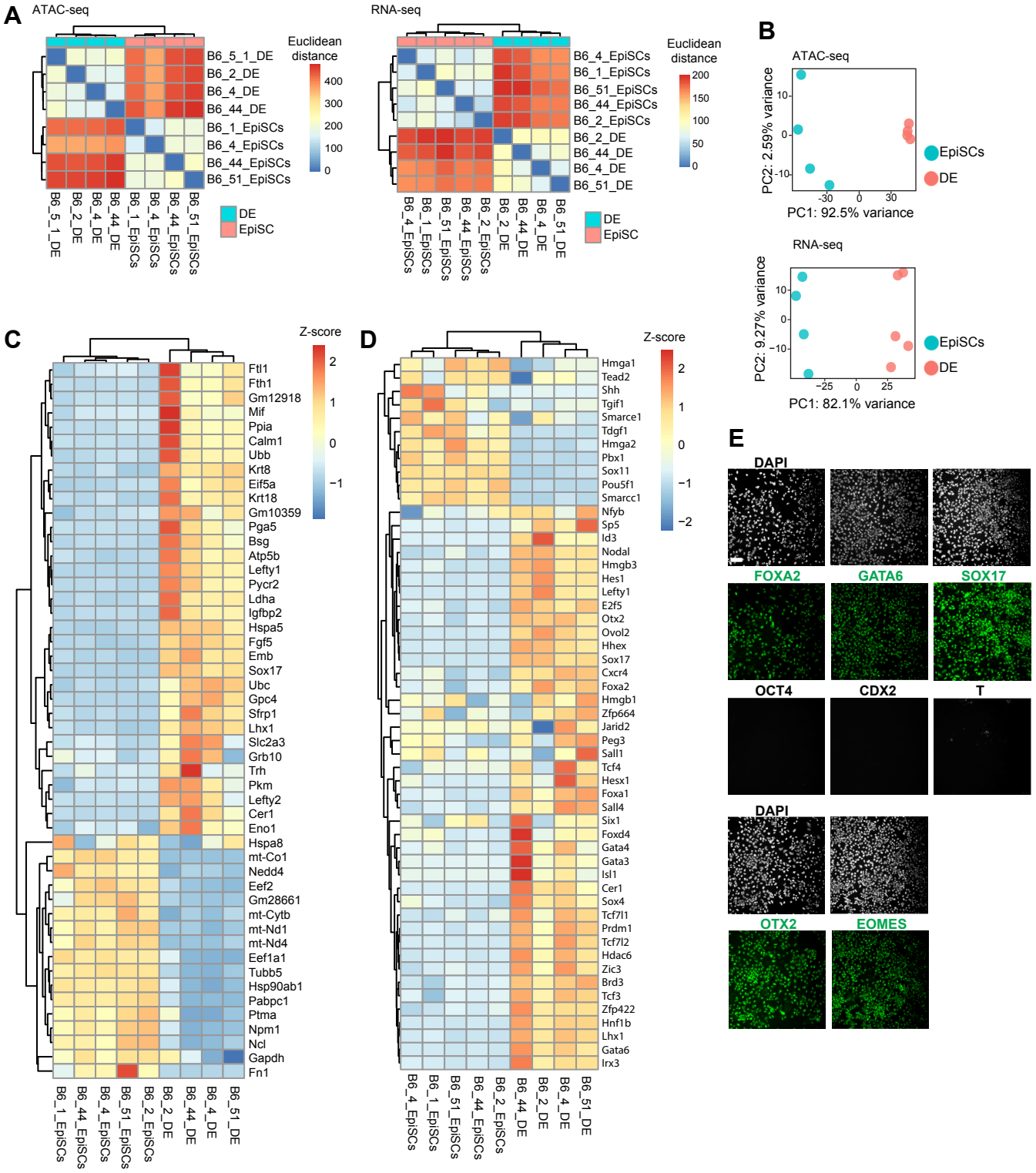


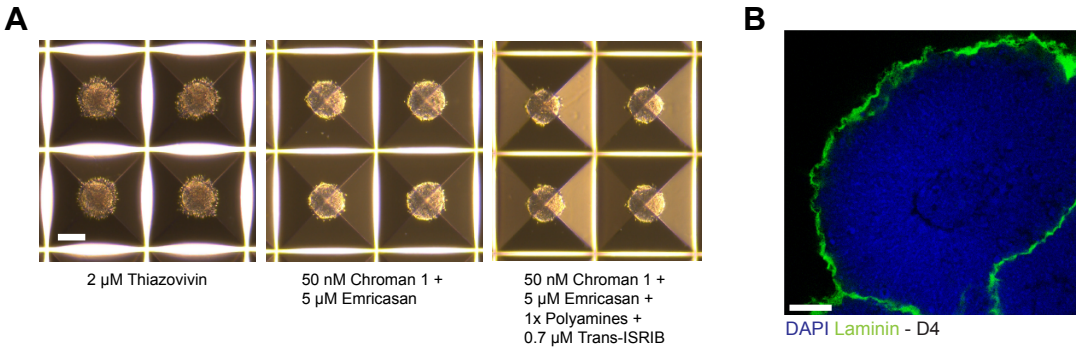
Figure 6: Characterization of day 12 forebrain organoids using scRNA-seq. (A) UMAP representation of scRNA-sequencing data from 5,652 cells isolated from day 12 organoids. Three broad classes of cell types are highlighted: oligodendrocyte progenitor cells (OPCs), hippocampus/cortical hem, and prethalamus. **(B)** Violin plots of gene expression levels for representative markers used to identify each cluster shown in **(A)** (see also **Supplementary table 5**). **(C)** UMAP colored based on single-cell expression of *Sox2*, a neuronal progenitor marker, or *Tubb3*, a post-mitotic neuron marker. **(D)** Cell cycle analyses on d12 scRNA-seq data. **(E)** UMAP representations of cells within putative prethalamic clusters only, highlighting expression of known pre-thalamic markers, from **(A). (F)** UMAP representations of gene expression data from only cells within putative hippocampal/cortical hem clusters from **(A).** We observed mutually exclusive expression of *Ctip2* (*Bcl11b*) and *Neurod1*, which delineate the early development of the CA1-2 and CA3 regions of the developing hippocampus, respectively (Simon et al., 2012). **(G)** UMAP representations of gene expression data from only cells within putative OPC clusters from **(A). (H)** VoxHunt analyses on the scRNA-seq dataset, separated by cluster, compared to the E18 developing mouse brain. Abbreviations: prosomeres 1–3 of the diencephalon (p1, p2, and p3), midbrain mesomere 1 and 2 (m1 and m2), and rhombomeres 1–11 (r1 to r11).



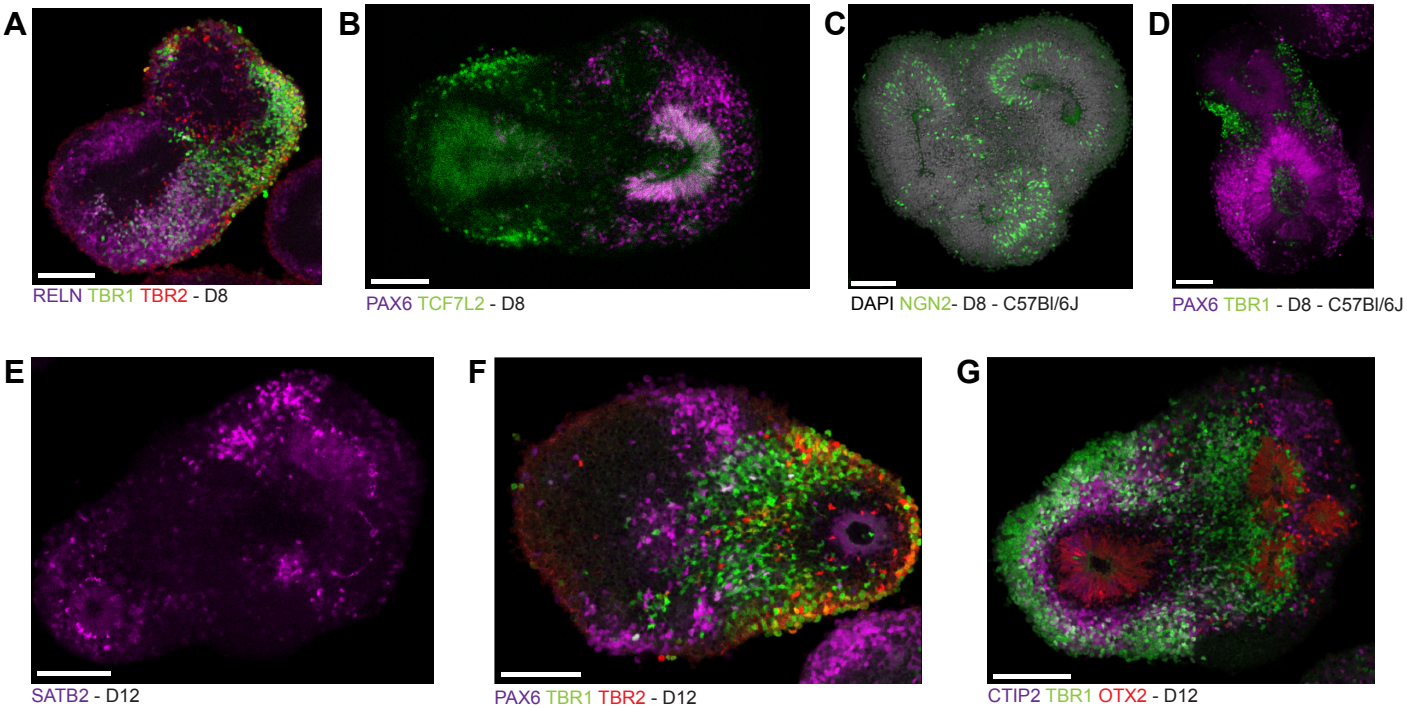
Supplementary Figure 1. Systematic optimization of conditions for EpiSC-DE differentiation. (A) Overview of state-of-the-art differentiation protocols for human (purple) and mouse (green) PSCs. Further details of each study are provided in Supplementary Table 1). **(B-C)** Results of DE protocol optimization experiments. For each condition, immunofluorescence staining was performed following 40 h of differentiation (end of Stage 2). Immunofluorescence images for the indicated antibodies were quantified using CellProfiler. Markers of DE (FOXA2) and mesoderm (CDX2) are shown. **(B)** Summary of results from anterior primitive streak (Stage 1) optimization experiments (n=2 technical replicates). After the indicated conditions for Stage 1, all conditions were exposed to the same Stage 2 conditions, and then quantified. **(C)** Summary of results for optimization of DE commitment (Stage 2) conditions (n=2 technical replicates). For Stage 1, the optimal conditions identified in **(B)** were used, and then cells were switched into the indicated conditions for Stage 2.



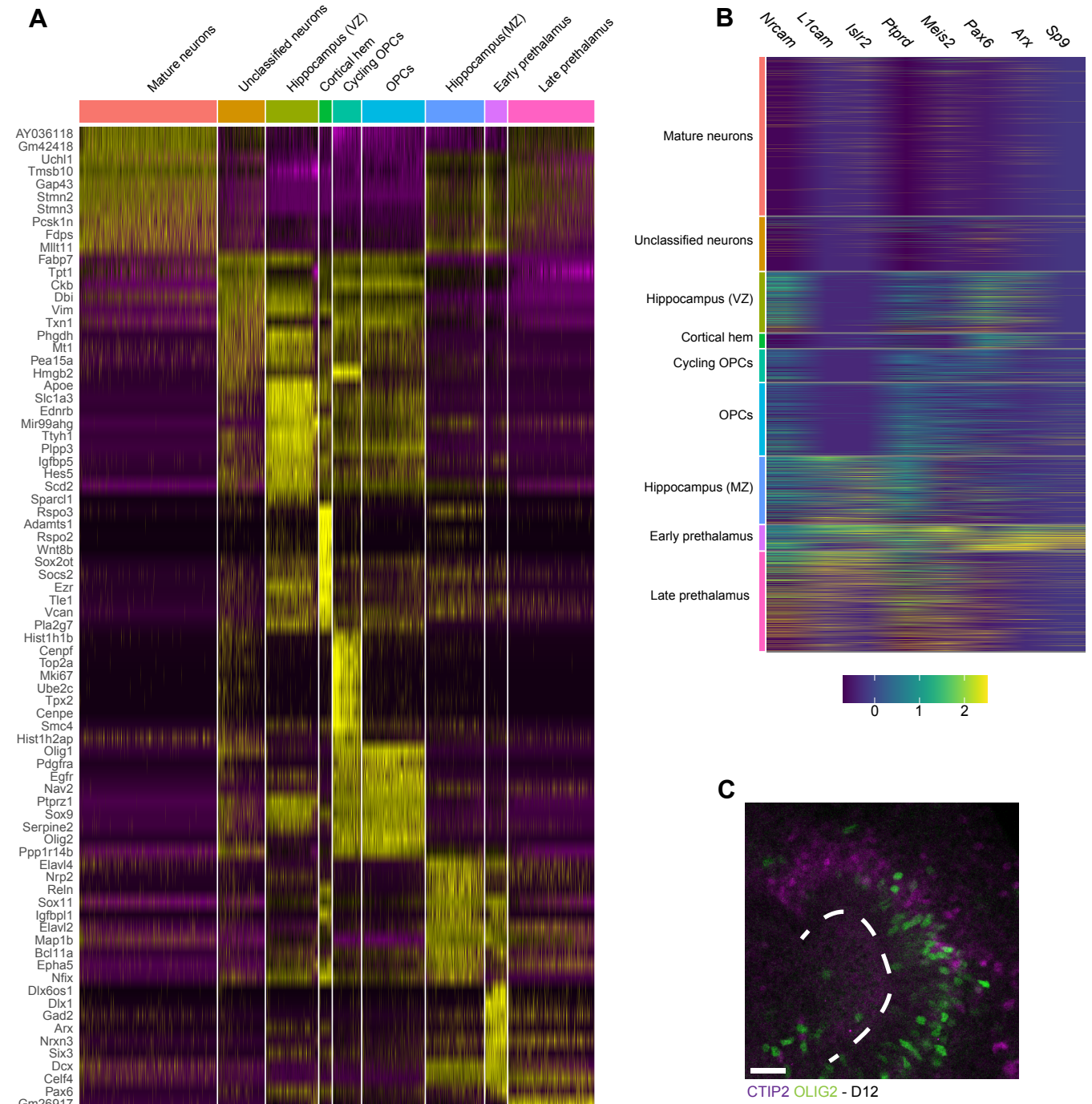
Supplementary Figure 2. Extended characterization of EpiSC-derived definitive endoderm. (A) Euclidean distance between samples based on ATAC-seq (n=4 per condition, left) and RNA-seq (n=5 for EpiSCs and n=4 for MACs-sorted DE samples, right). **(B)** PCA plot of bulk chromatin accessibility (left) and gene expression (right) of the same C57Bl/6J EpiSC and DE samples as **(A). (C)** Unbiased clustered heatmap of the top variable genes as detected by RNA-seq. **(D)** Expression of genes previously identified as DE markers in single cell gene expression studies of hPSC-derived DE and DE isolated from E7.0 embryos (Genga et al., 2019; Nowotschin et al., 2019). **(E)** Representative immunofluorescence images from DE differentiations for a panel of DE markers (FOXA2, GATA6, SOX17, OTX2, EOMES), pluripotency markers (OCT4), and primitive streak/mesoderm markers (T, CDX2). Scale bar = 50 µm.



Supplementary Figure 3: Optimization of conditions for forebrain organoid generation from mouse EpiSCs. (A) Brightfield images of d1 organoids in Aggrewell wells. EBs were formed in the presence of Thiazovivin or chroman 1 + emricasan (CE), or the recently published CEPT cocktail (chroman 1, emricasan, polyamine mix, trans-ISRIB) (Chen et al., 2021). Scale bar = 100 μm. **(B)** Confocal immunofluorescence image of a day 4 EB/organoid. Scale bar = 25 μm.



Supplementary Figure 4: EpiSC-derived brain organoids generate prethalamic and cortical-like neuronal populations. (A,B) Confocal immunofluorescence images of d8 organoids stained with antibodies against classical cortical markers (TBR1, TBR2, RELN and PAX6), which are also found in other regions of the brain such as the hippocampus, the cortical hem and the prethalamus, and thalamic markers (TCF7L2). Scale bar = 100 µm. **(C-D)** Confocal immunofluorescence of d8 organoids from C57Bl/6J background. Scale bar = 100 µm. **(E)** Confocal immunofluorescence image of a d12 organoid stained using antibodies against the classical upper-layer cortical marker SATB2. Scale bar = 100 µm. **(F-G)** Confocal immunofluorescence images of d12 organoids. Scale bar = 100 µm.



Supplementary Figure 5: scRNA-seq analyses on day 12 organoids reveals three main cell types. (A) Heatmap of the top 10 genes that define each cluster. Abbreviations: VZ = ventricular zone, MZ = mantle zone. **(B)** Expression profile, across clusters, of several prethalamic (*Sp9, Arx, Pax6, Meis2, Ptprd, Islr2*) and axonal (*L1cam, Nrcam*) markers. **(C)** Confocal immunofluorescence image of a day 12 organoid. Scale bar = 30 µm.