

# **NeMO Analytics: a compendium of transcriptomic data for the exploration of neocortical development**

---

In the format provided by the  
authors and unedited

## **.Supplemental Information**

**Supplemental Table 1:** Gene loadings and enrichment analyses related to the jointNMF() decomposition of scRNAseq data from mouse (La Manno 2021; 11) macaque (Micali 2023; 12), and human (Trevino 2021; 7) fetal neocortical tissue detailed in Figures 1 & 2 (p7CtxDev). In file “SonthaliaEtAl\_SupplTable1\_p7CtxDev.xlsx”

**Supplemental Table 2:** Gene loadings and enrichment analyses related to the jointNMF() decomposition of scRNAseq data from mouse (La Manno 2021; 11) macaque (Micali 2023; 12), and human (Trevino 2021; 7) fetal neocortical tissue detailed in Figure 3 (p40CtxDev). In file “SonthaliaEtAl\_SupplTable2\_p40CtxDev.xlsx”

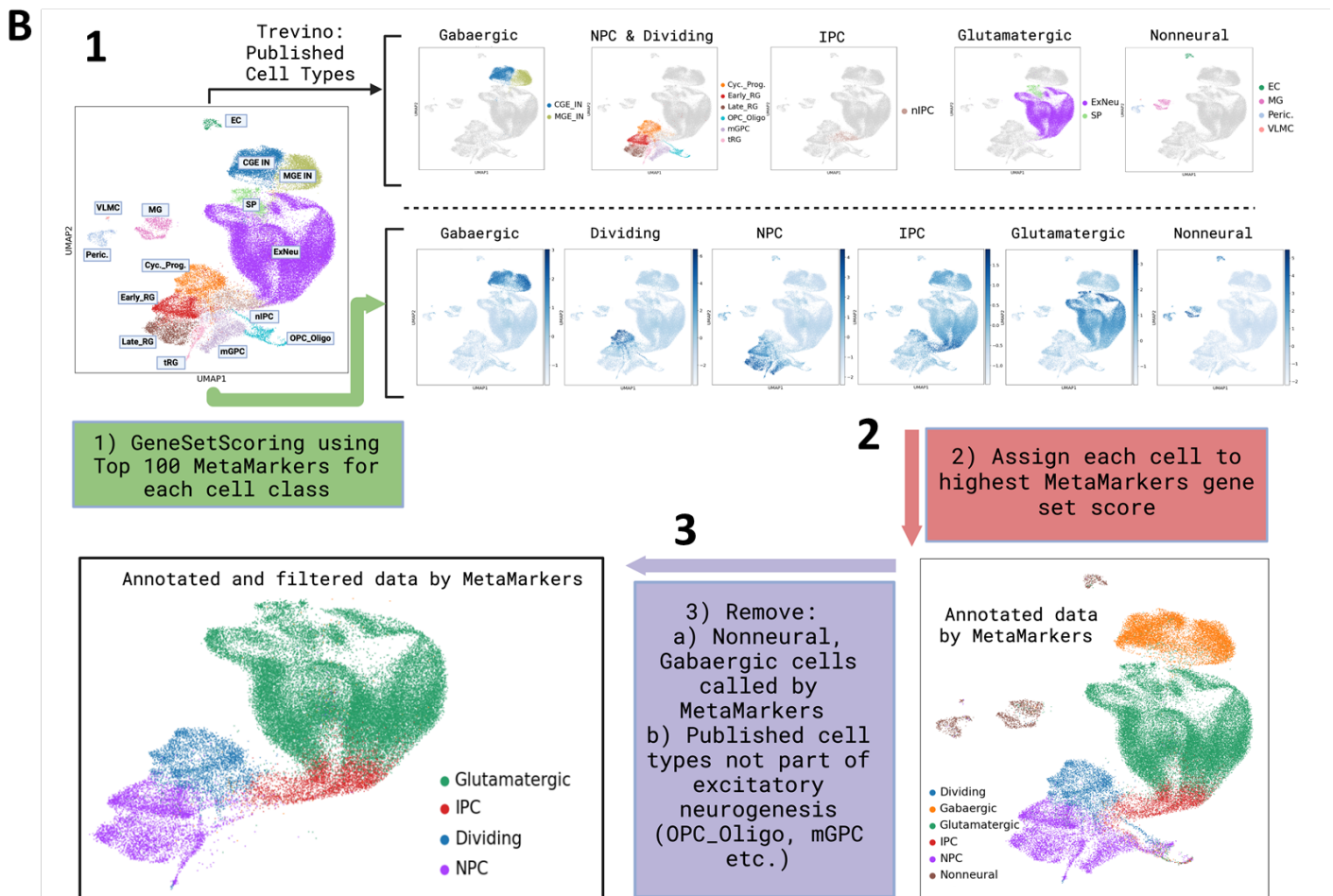
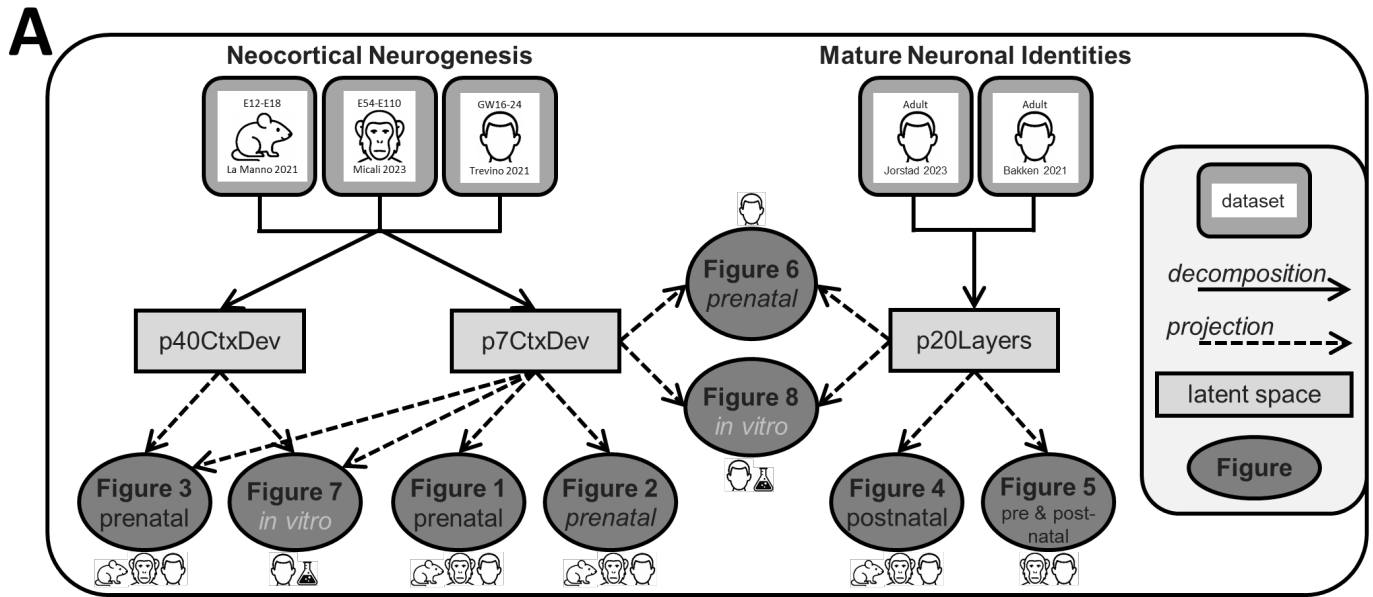
**Supplemental Table 3:** Gene loadings and enrichment analyses related to the jointNMF() decomposition of snRNA-seq (SMART-seq v4) data from adult human neocortical tissue (Jorstad 2023; 38 and Bakken 2021; 39) detailed in Figure 4 & 5 (p20CtxLayer). In file “SonthaliaEtAl\_SupplTable3\_p20CtxLayer.xlsx”

**Supplemental Figures 1-8:** [In this .pdf file]

**Supplemental References:** [In this .pdf file]

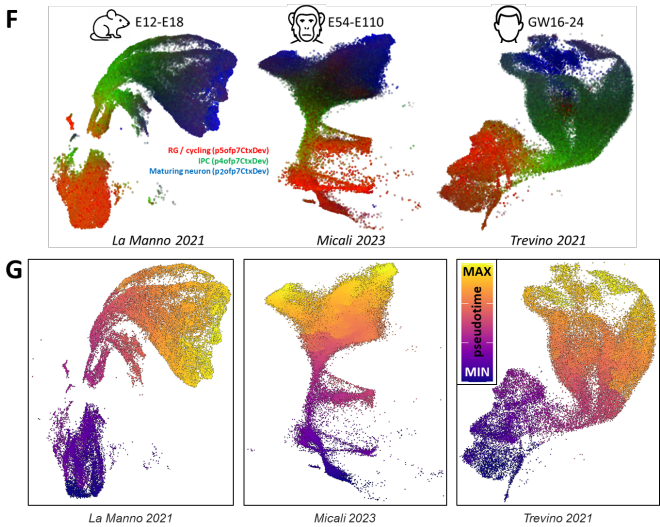
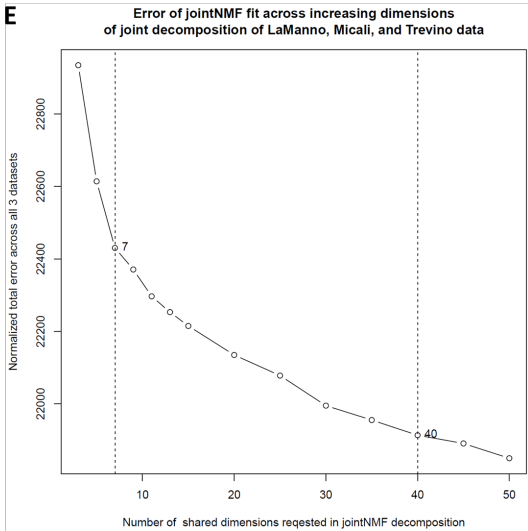
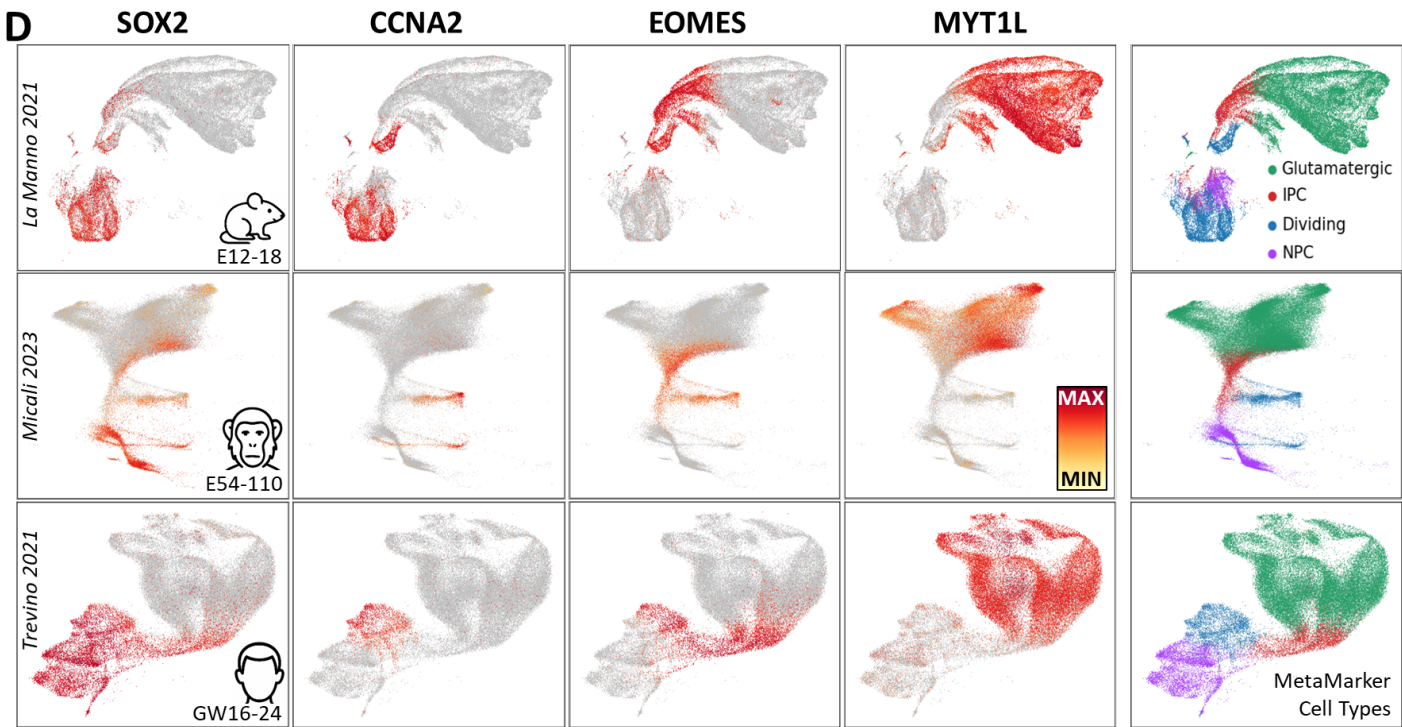


**Supplemental Figure 1 - related to main Figure 1**



C

|                        |                    | MetaMarker Consensus Cell Type Calls |          |       |               |           |           |       | %       |
|------------------------|--------------------|--------------------------------------|----------|-------|---------------|-----------|-----------|-------|---------|
| Author Cell Type Calls | La Manno 2021      | NPC                                  | Dividing | IPC   | Glutamatergic | Gabaergic | Nonneural | Other | Row Sum |
|                        | Radial glia        | 25.43                                | 67.48    | 1.32  | 0.01          | 0.37      | 5.29      | 0.09  | 99.9    |
|                        | Neuroblast         | 1.48                                 | 12.96    | 29.72 | 54.70         | 0.56      | 0.46      | 0.12  | 99.88   |
|                        | Neuron             | 0.01                                 | 0.07     | 0.42  | 99.36         | 0.13      | 0.01      | 0.01  | 100     |
|                        | Micali 2023        | NPC                                  | Dividing | IPC   | Glutamatergic | Gabaergic | Nonneural | Other | Row Sum |
|                        | Dorsal NSC         | 86.76                                | 13.17    | 0.00  | 0.03          | 0.00      | 0.05      | 0.00  | 100.01  |
|                        | enIPC              | 6.74                                 | 27.81    | 37.60 | 27.34         | 0.09      | 0.17      | 0.25  | 99.75   |
|                        | Excitatory Neuron  | 0.14                                 | 0.02     | 0.85  | 98.79         | 0.06      | 0.11      | 0.02  | 99.99   |
|                        | Trevino 2021       | NPC                                  | Dividing | IPC   | Glutamatergic | Gabaergic | Nonneural | Other | Row Sum |
|                        | Progenitor         | 90.22                                | 6.94     | 0.71  | 1.83          | 0.30      | 0.00      | 0.00  | 100     |
|                        | Cycling Progenitor | 4.08                                 | 93.03    | 2.41  | 0.37          | 0.11      | 0.00      | 0.00  | 100     |
|                        | nIPC               | 9.90                                 | 31.77    | 54.34 | 2.15          | 1.77      | 0.00      | 0.08  | 100.01  |
|                        | Excitatory Neuron  | 0.11                                 | 0.04     | 7.94  | 91.56         | 0.34      | 0.01      | 0.01  | 100.01  |



H

La Manno 2021

Micali 2023

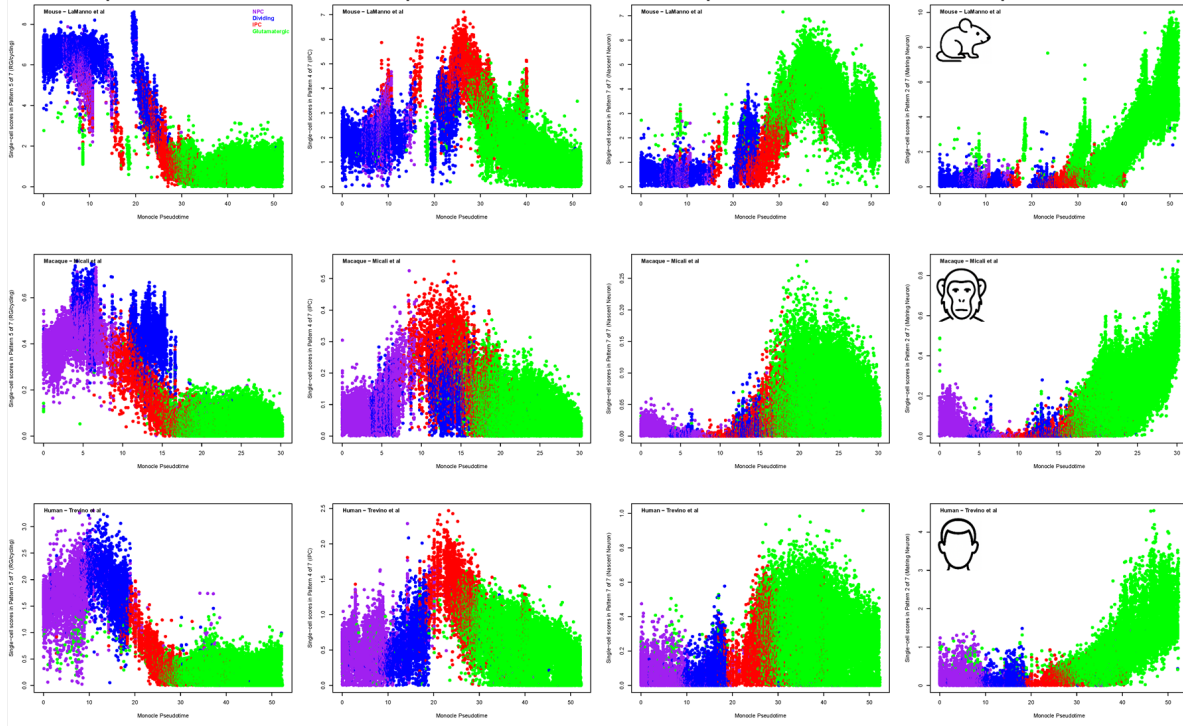
Trevino 2021

RG / cycling:  
p5of7CtxDev

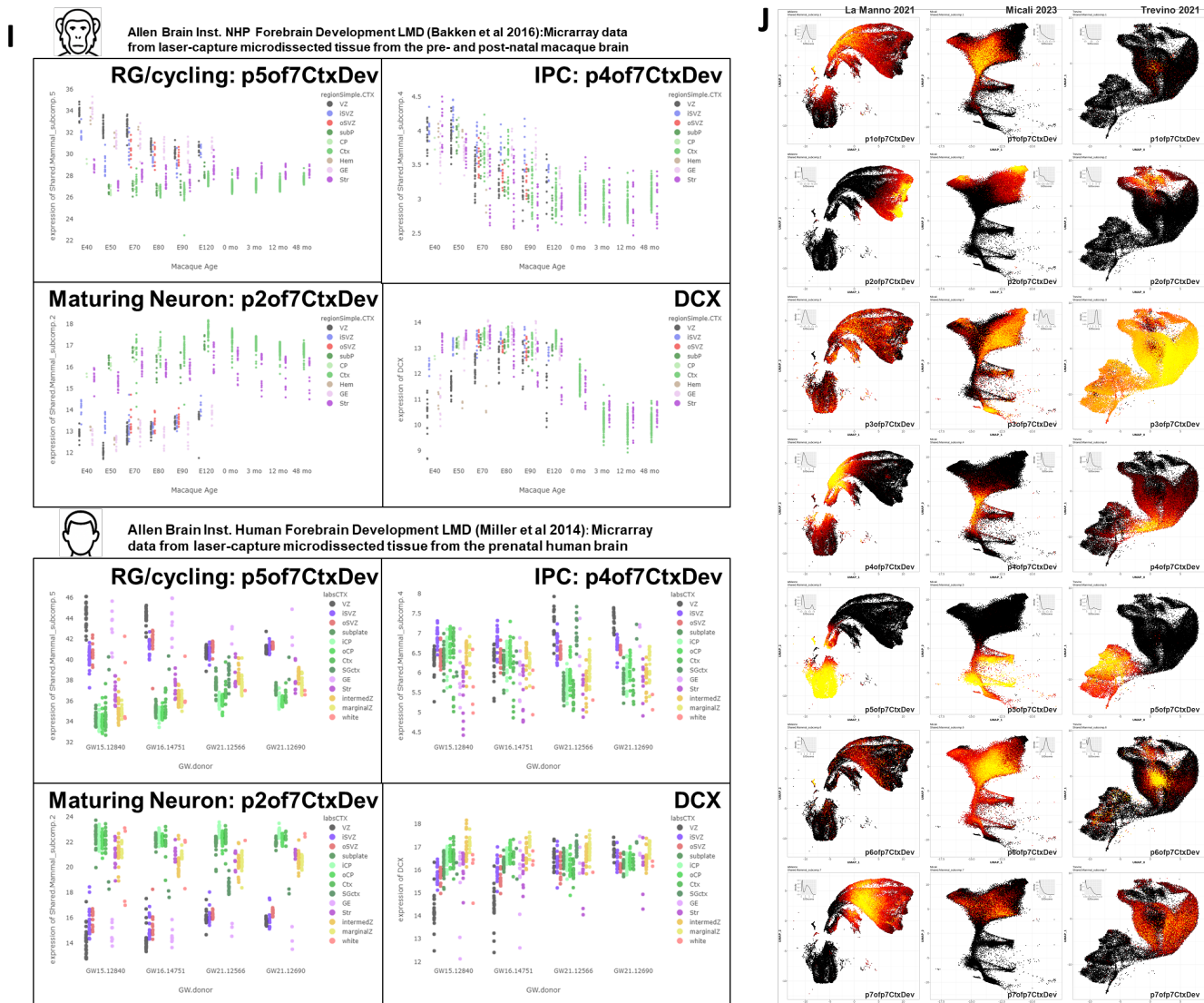
IPC:  
p4of7CtxDev

Nascent neurons:  
p7of7CtxDev

Maturing neurons:  
p2of7CtxDev







**Supplemental Figure 1 - related to main Figure 1: A]** Schematic of NeMO Analytics datasets used in decompositions and projections in this study **B]** Example of MetaMarker (24) classification procedure on Trevino 2021 dataset used for cell type calls in Figure 1 and 3. (1) Large panel and top row: Original author-published UMAP colored by author-curated cell types. Bottom row: Published UMAP colored by each of the 6 MetaMarker gene set scores calculated for each cell using the top 100 marker genes for each MetaMarker class (Methods). (2) Each cell was assigned to the MetaMarker class with the highest gene set score in that cell. (3) Nonneural and Gabaergic cells called by MetaMarkers classification were removed as well as author labeled cell-types not part of the excitatory neurogenic lineage. Further details regarding the procedure are in Methods and Code. **C]** We examined the correspondence between the original author cell type calls and the consensus MetaMarker cell type calls, summarizing them in this table. Values are colored by their percentage of calls falling into each category. As expected, of the 7 MetaMarker cell types, nearly no calls were made for 3 classes as these cell types had been omitted to explore excitatory neurogenesis (Gabaergic, Nonneural, and Other). Among the remaining 4 cell types (NPC, Dividing, IPC, Glutamatergic), there were highly variable levels of concordance: Excitatory neuron calls were >90% concordant in all three datasets. In the human Trevino data, both progenitor and dividing progenitor calls were also >90% concordant with MetaMarker cell type calls. However, among IPCs called in Trevino, 54% were called IPC and 32% were called dividing progenitors by MetaMarkers. Similarly, in Micali et al, author neuroblast cell types calls (a cell type designation that could be viewed as spanning both new neurons and IPCs) were assigned to 30% IPCs and 55% Glutamatergic. These large discordances in cell type calls are understandable, e.g. any NPC or IPC is an NPC or IPC, and can be cycling or not as well. There are likely few objectively incorrect cell type calls being made. These inconsistencies reflect the insufficiency of individual, non-overlapping cell class labels in capturing cell types and states, especially in dynamic developmental contexts. **D]** Colorized UMAP plots for each of the three datasets used in Figure 1 showing the expression of 4 individual genes representative of broad cell types in neurogenesis: SOX2 in neural progenitors, CCNA2 in cycling progenitors, EOMES in intermediate progenitors, and MYT1L in new neurons. UMUPS in the right most column are colored by the MetaMarker cell type calls as made in panel A. **E]** To determine a rational number of dimensions for each joint decomposition, we examined the change in modeling error as we increased

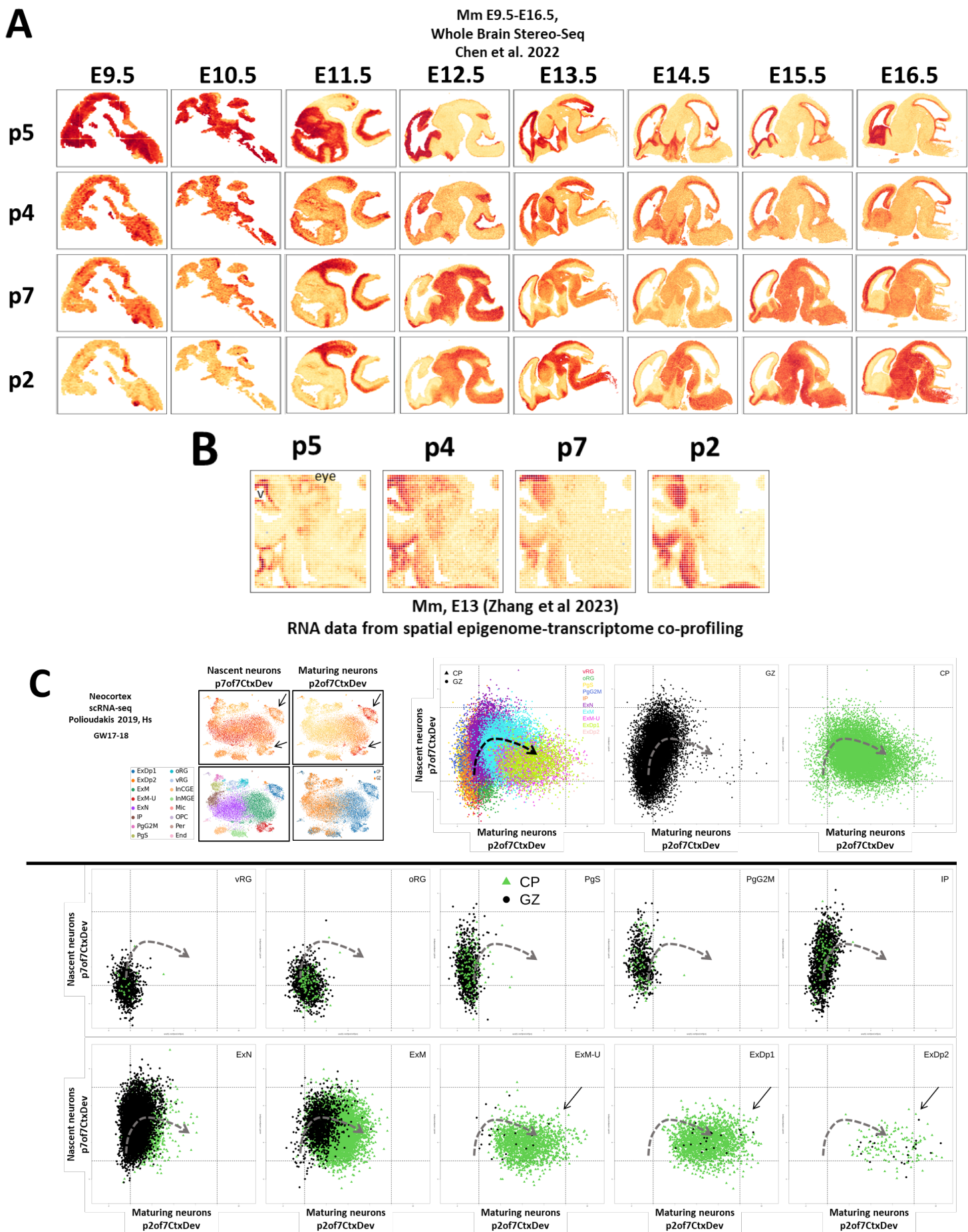
the number of requested dimensions. Error at 7 and 40 (the number of dimensions selected for the Figure 1 and Figure 3 analyses, respectively) are highlighted with vertical dashed lines. **F]** To illustrate the continuous and overlapping nature of the signals defined in joint decomposition, we have colored each of the three datasets in this analysis by 3 of the 7 decomposed transcriptomic programs, mixing colors where individual cells express more than 1 program. **G]** Same three datasets colored by Monocle 3 pseudotime. **H]** Plotting of individual single-cell scores from each of the highlighted jointNMF signatures (Y-axis) across pseudotime calculated with Monocle 3 (X-axis), colored by the consensus MetaMarker cell type calls. **I]** In Figure 1D we see that the RG/cycling pattern (p5of7CtxDev) progressively declines in progenitors, while the maturing neuron pattern (p2of7CtxDev) increases in progenitors over time and the IPC pattern (p4of7CtxDev) is reduced over time in neurons. Projection of microarray data from laser capture microdissected (LMD) regions of the developing macaque and human brain into these transcriptome dimensions confirm all of these effects. While these effects could have been related to “ambient” RNA effects from unintentionally lysed cells in the scRNA-seq data, their confirmation in the LMD data demonstrates that this is not the case ([NeMOLinkS01](#)). Rising expression of DCX in the VZ in these same data is a single gene example of the “neuronization” of progenitors and can also be viewed online at [NeMOLinkS02](#). Despite the fact that our joint decomposition focused on mid-gestation, the conserved transcriptomic elements of neurogenesis in the neocortex are clearly observable in the 1st trimester Braun data (19) where the expected cell types express each of the fundamental components of neurogenesis. Our transcriptomic patterns can be explored in the Braun data (among many other datasets) at: [NeMOLinkS03](#). **J]** All 7 individual patterns in the p7CtxDev jointNMF decomposition. Individual gene weights underlying these patterns and their gene set enrichments are in Table S2 and all 40 patterns can be explored across mammalian neocortical developmental data at [NeMOLink07](#).

### Notes on gene set enrichments in main Figure 1:

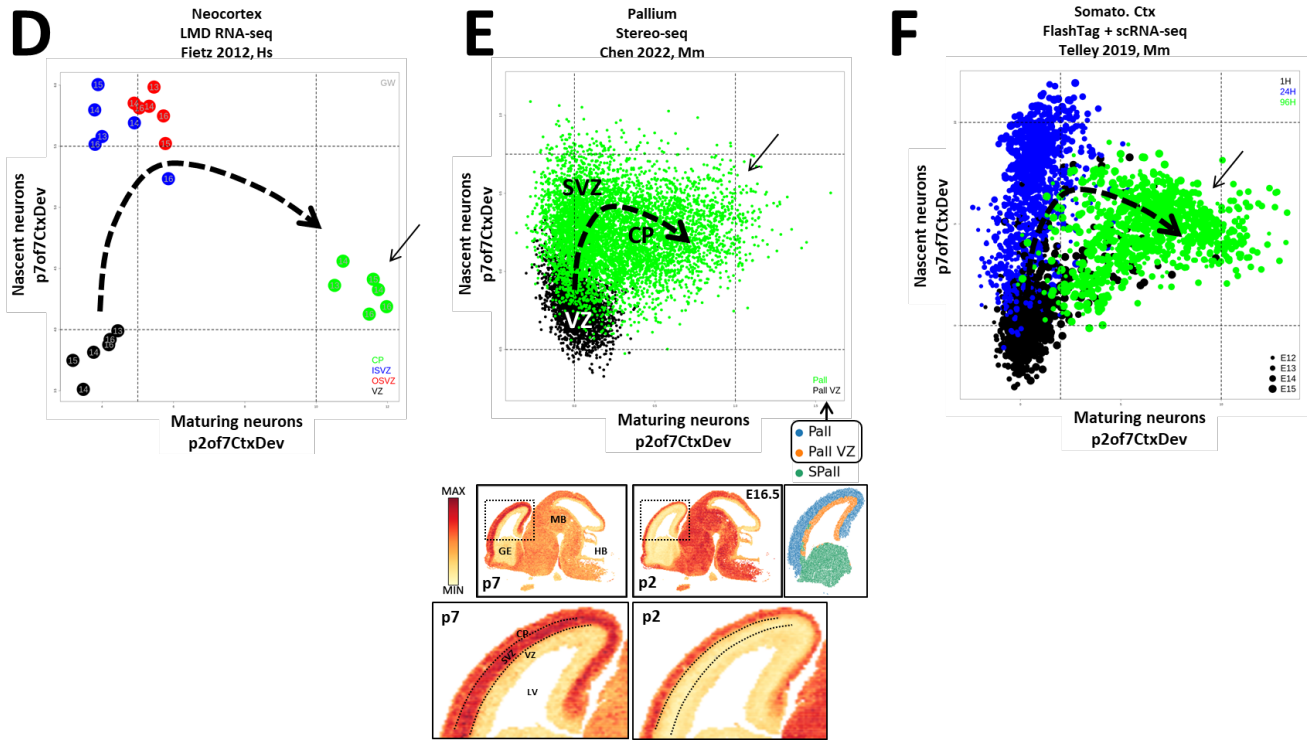
We interrogated the conserved transcriptomic elements of neocortical neurogenesis for association with 1] cellular processes, 2] high penetrance, low frequency disease variant gene lists, and 3] polygenic risk using GWAS data and MAGMA [polygenic risk known to play central roles in complex human brain disease using gene level summaries of recent genome-wide association studies in neuropsychiatric disease and brain structure] (Figure 1C, Methods; GSEA code can be found on our [GitHub](#) site). Consistent with many recent observations of their broad polygenic nature (112, 113, 114), the analysis of genome-wide risk yielded strongest associations for schizophrenia and bipolar disorder with the neuronal transcriptome patterns p7 and p2of7CtxDev (Figure 1C). Interestingly, the later neuronal pattern p2 also showed significant association with cortical structure phenotypes. To explore high penetrance, low frequency genetic variation linked to brain disease in these patterns, we conducted enrichment analysis on gene loadings using lists of genes discovered in genome sequencing studies of disease. This revealed that diseases which disrupt the gross structure of the brain, including microcephaly (115) and hydrocephaly (116, 117, 118, 119, 120), are associated with p5 that is high in radial glial cells (RGCs) and especially in cycling progenitors of the developing cortex, consistent with additional recent observations (26). Genes harboring high penetrance, low frequency variants linked to neuropsychiatric and neurodevelopmental disorders were strongly enriched in the neuronal patterns p7 and p2, with ASD having particularly strong associations. This is consistent with many observations indicating that low-frequency, high-penetrance *de novo* variants play a key role in this disorder (121, 122).

While the neuronal transcriptome patterns most enriched in synapse associated genes (p7 and p2of7CtxDev) show the highest enrichments for ASD, SCHZ, and BD (Figure 1C, consistent with 123, 75, 124, 125), we must take caution in interpreting these enrichment analyses that have become ubiquitous in efforts to use gene expression data combined with GWAS results in order to infer mechanisms of GWAS variants. The basic premise is sound: regulatory GWAS variants can only function to confer risk for disease in cells where they affect gene expression, so we should look at cell types where genes linked to GWAS variants are expressed. This thinking would for instance indicate that we should look to brain rather than liver for the function of ASD, SCHZ, and BD associated variants because affected genes are expressed in brain and not liver. The reasoning breaks down when comparing multiple tissues or cell types which express the risk-associated genes at different levels. There is no more reason to implicate cells that have higher expression rather than lower. As long as the genes are expressed in a cell or tissue, there is a possible functional role. So focus should be on the number of risk genes expressed in a cell type or tissue rather than the level at which it is expressed, but virtually all conventional enrichment testing assesses expression levels. Additionally, given that these are developmental disorders, a focus on the cells where the genes are first expressed may be more effective than where genes are expressed at the highest levels. Even if levels of risk genes are lower in neural progenitors than neurons, it may be there that risk is likely to play out. This is consistent with recent studies showing that genes previously considered to be solely synaptic in function also have much earlier roles in neural progenitor function (126, 127). Adding to the notion that earlier events may harbor risk mechanisms, a recent study has indicated that genetic effects on gene expression (eQTLs) are strongest during *in utero* brain development before the formation of most neocortical synapses (128). Another of our recent reports has examined risk in progenitors in detail (26).

Supplemental Figure 2 - related to main Figure 2





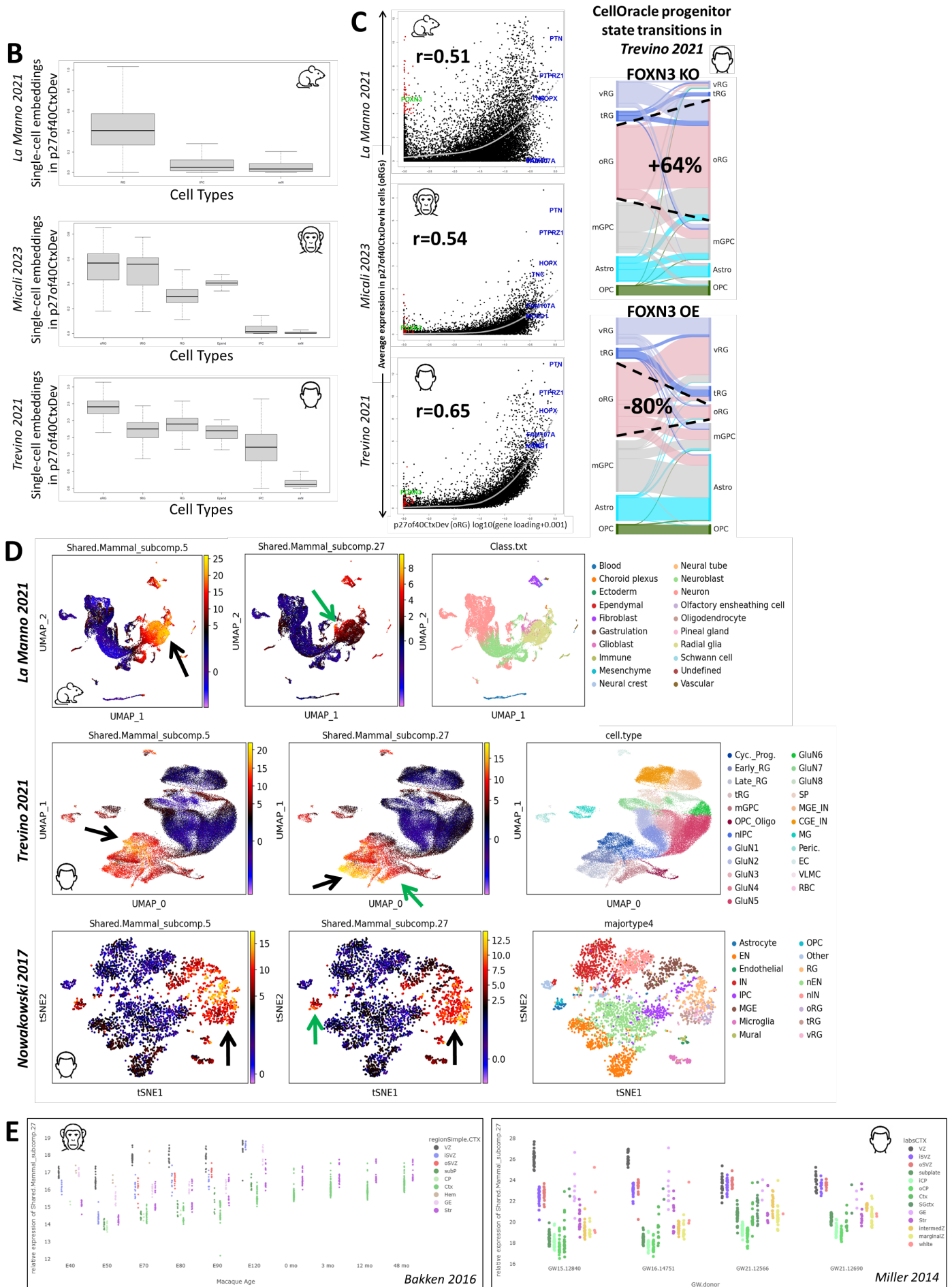


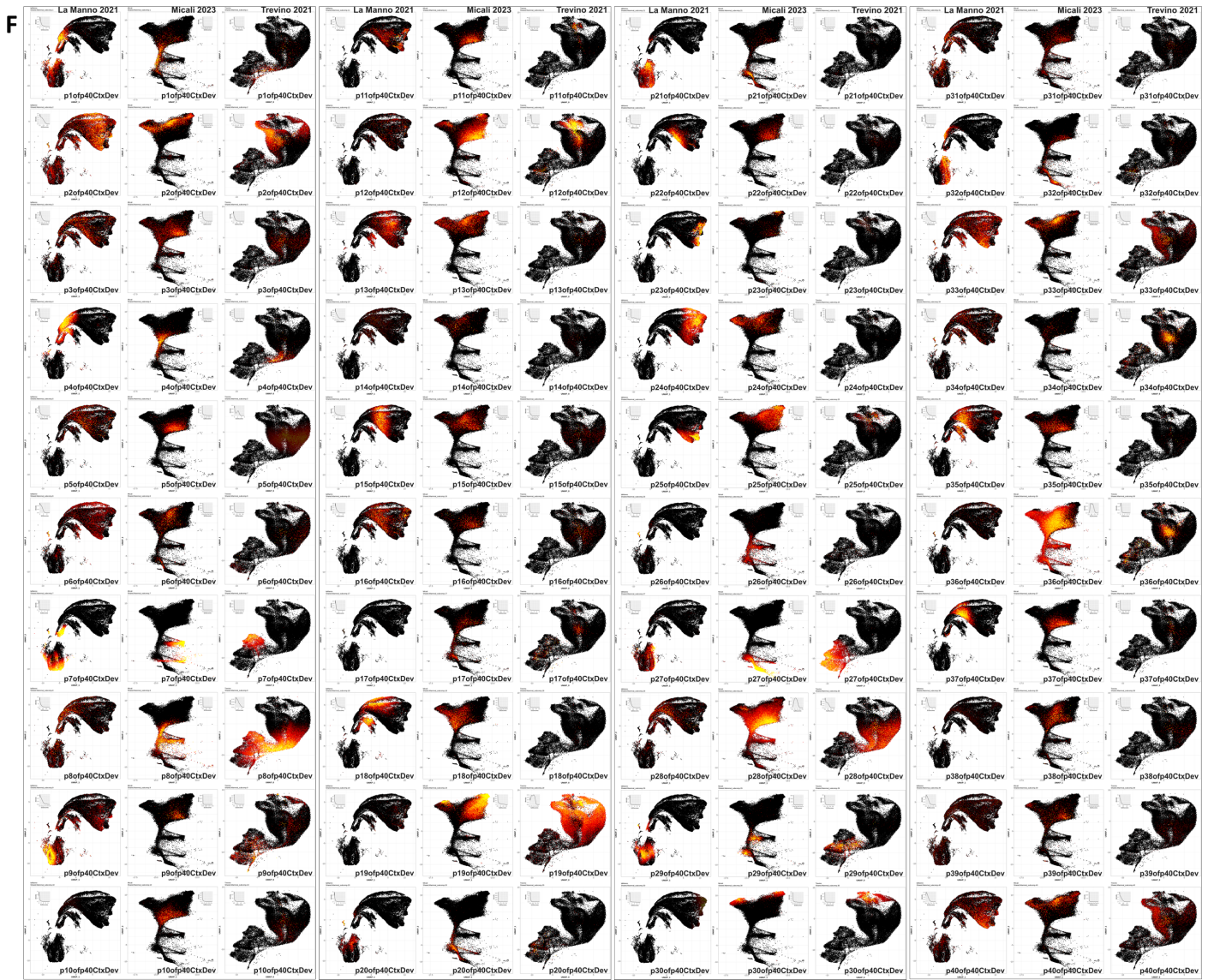
**Supplemental Figure 2 - related to main Figure 2:** **A]** Visualization of projected embeddings in p5, p4, p7 and p2of7CtXDevo for individual spatial measurements in full sagittal brain slices from E9.5-E16.5 in the Chen 2022 dataset (54). These sections indicate clearly that both the progenitor and neuronal transcriptome signatures are not specific to the neocortex, but rather are pan-brain signatures of these broad cell types. Many spatio-temporal details can be seen across these sections as the developing brain progresses from a tissue dominated by progenitors to a tissue of post-mitotic neurons. **B]** Projected embeddings in p5, p4, p7 and p2of7CtXDevo for individual RNA measurements from spatial epigenome-transcriptome co-profiling (42). Here the outward radial progression of cells expressing p5 (RG at the ventricle), to p4 (delaminated IPCs), to p7 (nascent neurons) and finally p2 (maturing neurons) can be seen. V=ventricle, lateral, of the forebrain. **C]** Related to data from Polioudakis 2019 (28) in Figure 2: Upper left panels show the original author published tSNE with author's cell type calls and plots colored by p7 (proneural nascent neuron signature) and p2of7CtXDevo (maturing neuron signature). Original author cell type calls: Ex=excitatory neurons, Dp=deep, N=new/migrating, M=maturing, Ip=intermediate progenitor, Pg=cycling progenitor in S or G2M phase, RG=ventricular (v) or outer (o) radial glia, In=inhibitory neurons of the medial (MGE) or caudal (CGE) ganglionic eminence, Mic=microglia, OPC=oligodendrocyte precursor cell, Per=pericyte, End=endothelial cell. Three panels at upper right show projected embeddings of cells in the excitatory neurogenic trajectory (according to author cell type calls) as they move through the transcriptomic dimensions defined by p7 and p2of7CtXDevo, first colored by individual cell types and then separately by where the single cells were dissected from in the developing cortex. GZ=germinal zones, CP=cortical plate (both these dissections included portions of the intermediate zone – this is especially relevant to panel G in main Figure 2, where the intermediate zone was dissected separately in the macaque developing neocortex and show the descent of p7 as p2of7CtXDevo begins to rise). Lower two rows of panels show individual cell types in the p7 v p2 transcriptomic space (colored by their dissection of origin) as they progress through neurogenesis. Throughout panel A, vertical and horizontal dashed lines are in identical positions in each plot to enable the comparison of p7 and p2of7CtXDevo pattern values across cell types in different plots. **D]** Scatter plots of projected embeddings in p7 and p2of7CtXDevo for individual samples from LMD coupled RNA-seq in the developing human neocortex from the Fietz 2012 (53) dataset. LMD=laser microdissection. CP=cortical plate, OSVZ=outer subventricular zone, ISVZ=inner subventricular zone, VZ=ventricular zone. **E]** Scatter plot of projected embeddings in p7 and p2of7CtXDevo for individual spatial measurements in the Chen 2022 dataset (54). The spatial segmentation of pallium (Pall) and pallium ventricular zone (Pall VZ) used by the original authors is used here to color these individual spatial measures in the scatterplot. These same projected embeddings are also shown in NeMO across the spatial section beneath the scatter plot. **F]** Scatter plot of projected embeddings in p7 and p2of7CtXDevo for Flash-Tag labeled (birth-dated) single-cells in the Telley 2019 dataset (27). The time from terminal division is represented by color in the scatter plot. In panels **C-F**, arrows indicate maturing neurons, where p7of7CtXDevo has descended and p2 is highest. Along with main Figure 1G, this temporal and spatial understanding of new-born neurons as they move through these transcriptomic dimensions is central for their use in Figures 6 and 8 where we harness these same signatures to explore the protracted maturational trajectories of adult neuronal signatures.

**A**

|                        | G2M<br>p7of40CtDev                                                                                                         | S-phase<br>p29of40CtDev                                                                                      | oRG<br>p27of40CtDev                                                                                                       | p11of40CtDev                                                                                                                 | p19of40CtDev                                                                                                               | p25of40CtDev                                                                                                                                                                   | p30of40CtDev                                                                                                                                                          | Neuronal patterns<br>MetaMarker Cell Types |
|------------------------|----------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------|
| La Manno 2021          |                                                                                                                            |                                                                                                              |                                                                                                                           |                                                                                                                              |                                                                                                                            |                                                                                                                                                                                |                                                                                                                                                                       |                                            |
| Micali 2023            |                                                                                                                            |                                                                                                              |                                                                                                                           |                                                                                                                              |                                                                                                                            |                                                                                                                                                                                |                                                                                                                                                                       |                                            |
| Trevino 2021           |                                                                                                                            |                                                                                                              |                                                                                                                           |                                                                                                                              |                                                                                                                            |                                                                                                                                                                                |                                                                                                                                                                       |                                            |
| top genes              | TOP2A, CENPF, CDK1, HMGB2, CCNB1, CCNA2, PRCL1, CDC20, CENPE, CDCA3, KIF22, MKI67, CENPA, CCNB2, AURKA, AURKB, KIF23, BUB3 | MCM6, HELLS, MCM3, PCNA, MCM2, MCM5, GINS2, GMNN, CDCA7, MCM4, RFC3, DNAI9, RFC4, GAS1, TUBA1B, DACH1, HSPAS | VIM, PTN, TTYH1, PTPRZ1, ID4, SOX2, HOPX, DBI, SFRP1, HES1, BCAN, CNN3, FABP5, FOS, TMEM132B, HMGN3, HMGCS1, FAM107A, JUN | IGFBPL1, DCC, NFIX, PDE4D, NFIB, NTRK3, CACNA1C, CTNNA2, NAV3, SETBP1, AUTS2, TCF4, ANK3, CELF2, BACH2, EPHA4, KCNN2, BCL11A | NKTR, WSB1, ARID1B, ZEB2, SEMA3C, MAP4K4, SRGAP1, PCLO, CAMK2N1, MAP6, ARHGAP21, CELF1, CDK5R1, MYT1L, CCDC82, NCAM1, SYTA | SATB2, PTN, MEZF2, DAB1, DLG2, ROBO2, CACNA2D1, PTPRK, SYTA, PTPRD, GPMMA, STMN2, MAPT, STMN4, MYT1L, PTPRZ1, CELF2, NCAM1, MAP1B, NAV3, STMN3, KCNGB, RBFOX1, HMGCS1, NEUROD6 | MEF2C, NRCAM, DLG2, NR3G3, NTRK2, LMO3, SOX5, CELF4, GRIN2B, GRM5, PCDH9, FRICKLE1, RGS6, CACNAIE, DLGAP1, PRICKLE2, SCN2A, CAMKV, TLE4, PTPRN2, PCDH7, CAMK2B, FGF12 |                                            |
| GO                     | Cell Cycle p=2.4e-149<br>Synapse p=9.6e-1                                                                                  | Cell Cycle p=4.5e-60<br>Synapse p=5.6e-1                                                                     | Cell Cycle p=5.2e-34<br>Synapse p=1.4e-1                                                                                  | Cell Cycle p=6.3e-17<br>Synapse p=6.7e-22                                                                                    | Cell Cycle p=2.1e-21<br>Synapse p=3.2e-26                                                                                  | Cell Cycle p=3.6e-6<br>Synapse p=2.2e-54                                                                                                                                       | Cell Cycle p=1.3e-2<br>Synapse p=3.0e-102                                                                                                                             |                                            |
| High Pen.<br>Low Freq. | Microcephaly p=4.9e-16<br>Hydrocephaly p=8.2e-7<br>SFARI (ASD) p=1.8e-8<br>SCHEMA (SCHZ) p=1.1e-1                          | Microcephaly p=1.1e-4<br>Hydrocephaly p=1.3e-3<br>SFARI (ASD) p=2.5e-3<br>SCHEMA (SCHZ) p=1.2e-1             | Microcephaly p=1.1e-4<br>Hydrocephaly p=7.5e-6<br>SFARI (ASD) p=5.3e-3<br>SCHEMA (SCHZ) p=1.1e-1                          | Microcephaly p=2.0e-3<br>Hydrocephaly p=1.0e-3<br>SFARI (ASD) p=1.7e-38<br>SCHEMA (SCHZ) p=1.4e-7                            | Microcephaly p=1.1e-2<br>Hydrocephaly p=5.9e-3<br>SFARI (ASD) p=7.1e-34<br>SCHEMA (SCHZ) p=3.2e-9                          | Microcephaly p=4.5e-2<br>Hydrocephaly p=1.6e-2<br>SFARI (ASD) p=3.6e-40<br>SCHEMA (SCHZ) p=1.7e-7                                                                              | Microcephaly p=3.6e-1<br>Hydrocephaly p=1.1e-1<br>SFARI (ASD) p=2.6e-53<br>SCHEMA (SCHZ) p=1.5e-9                                                                     |                                            |
| Polygenic<br>(MAGMA)   | ASD p=2.9e-1<br>SCHZ p=6.8e-3<br>BD p=1.8e-3<br>Total surface p=3.9e-6<br>Global thickness p=2.7e-4                        | ASD p=2.6e-1<br>SCHZ p=2.2e-3<br>BD p=3.2e-2<br>Total surface p=1.8e-2<br>Global thickness p=3.9e-2          | ASD p=7.4e-2<br>SCHZ p=1.5e-4<br>BD p=5.0e-2<br>Total surface p=1.4e-1<br>Global thickness p=4.2e-1                       | ASD p=5.2e-6<br>SCHZ p=3.7e-18<br>BD p=2.2e-12<br>Total surface p=6.6e-7<br>Global thickness p=1.9e-6                        | ASD p=6.6e-5<br>SCHZ p=2.5e-22<br>BD p=8.2e-12<br>Total surface p=6.0e-7<br>Global thickness p=1.7e-8                      | ASD p=9.8e-5<br>SCHZ p=2.8e-25<br>BD p=7.6e-18<br>Total surface p=1.4e-2<br>Global thickness p=1.3e-4                                                                          | ASD p=2.6e-2<br>SCHZ p=1.6e-21<br>BD p=9.4e-14<br>Total surface p=4.9e-2<br>Global thickness p=1.3e-5                                                                 |                                            |
| La Manno 2021          |                                                                                                                            |                                                                                                              |                                                                                                                           |                                                                                                                              |                                                                                                                            |                                                                                                                                                                                |                                                                                                                                                                       |                                            |
| Micali 2023            |                                                                                                                            |                                                                                                              |                                                                                                                           |                                                                                                                              |                                                                                                                            |                                                                                                                                                                                |                                                                                                                                                                       |                                            |
| Trevino 2021           |                                                                                                                            |                                                                                                              |                                                                                                                           |                                                                                                                              |                                                                                                                            |                                                                                                                                                                                |                                                                                                                                                                       |                                            |
|                        | G2M<br>p7of40CtDev                                                                                                         | S-phase<br>p29of40CtDev                                                                                      | oRG<br>p27of40CtDev                                                                                                       | p11of40CtDev                                                                                                                 | p19of40CtDev                                                                                                               | p25of40CtDev                                                                                                                                                                   | p30of40CtDev                                                                                                                                                          | Neuronal patterns                          |





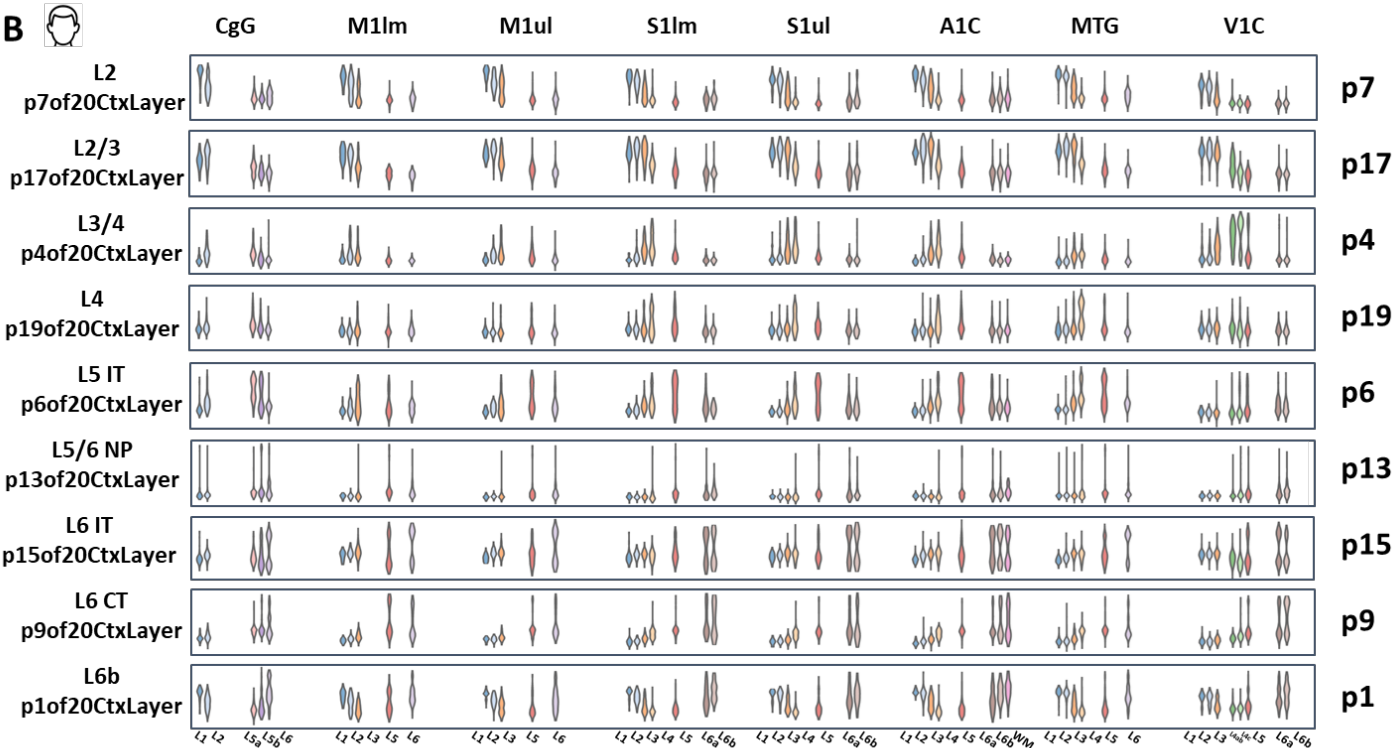
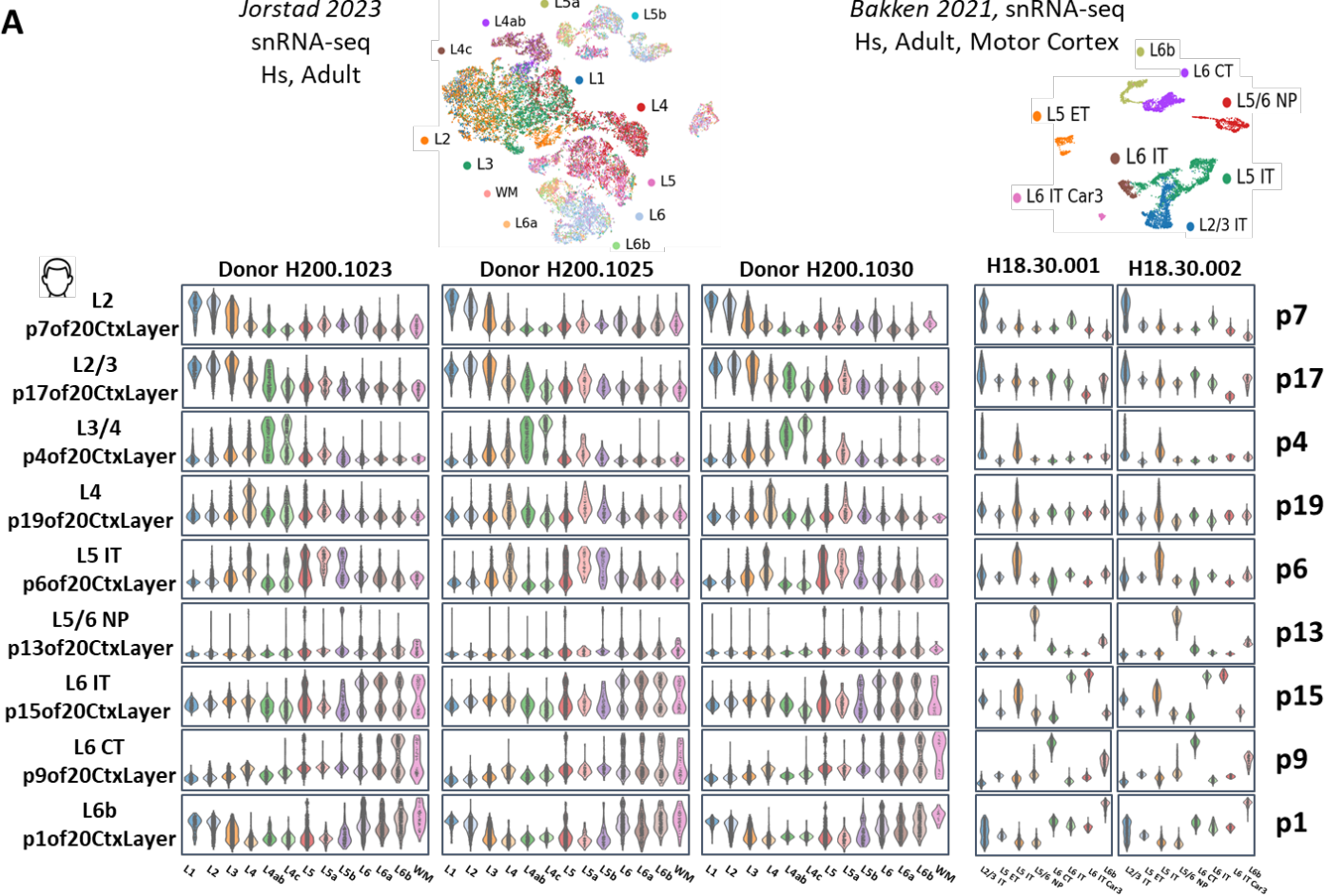


**Supplemental Figure 3 - related to main Figure 3: A]** Single-cell embedding UMAP plots of all three mammalian neocortical datasets for 7 selected patterns (4 patterns that appear in differing subsets of neurons and 3 that define transcriptomic programs within progenitors) from the full set of 40 (p40CtDev) along with individual genes and gene set enrichments for each. Many details can be seen in this higher resolution decomposition of the mammalian neocortical neurogenesis datasets. Somewhat paradoxically, gene loadings for neuronal patterns p11of40CtDev and p19of40CtDev show strong enrichments for both synaptic genes and genes involved in the cell cycle (Figure S3A). This again highlights the fact that transcriptomic programs span cell populations that are often defined as discrete non-overlapping classes, but which in fact share many molecular attributes. Interestingly, in addition to highly significant enrichments for neuropsychiatric diseases, enrichments in the genetics of brain structure are greatest in these patterns that appear to bridge proliferative and post-mitotic states in neurogenesis. Patterns p25of40CtDev and p30of40CtDev have much less cell cycle involvement and show even stronger enrichments in synaptic as well as disease associated gene sets (Figure S3A). Individual genes with especially high loadings in these patterns include well-known neuropsychiatric risk genes, including RBFOX1 and MEF2C, many synaptic genes, and layer-defining transcription factors (TFs), including SATB2 and TLE4. Hence, neuronal patterns defined in this higher resolution decomposition show enrichment in genes linked to human brain structure and disease, with hits of similar significance to those found in the low resolution decomposition (Figure 1), but here with greater precision in defining the subpopulations of genes and cells that are involved with these associations. Also in this higher resolution decomposition, transcriptomic elements of neural progenitors and the cell cycle have been separated in greater detail. Two new patterns which span distinct subpopulations of both RGs and IPCs are shown in Figure S3. Both the ranking of individual marker genes among the patterns' gene loadings (G2M marker MKI67 ranked #22 in p7of40CtDev and S-phase marker PCNA ranked #5 in p29of40CtDev) and enrichment analysis (G2M checkpoint in p7of40CtDev,  $p=4.9\text{e-}76$ ; DNA replication in p29of40CtDev,  $p=2.1\text{e-}19$ ) clearly indicate that these patterns are present specifically in the transcriptomes of cells in G2M (p7of40CtDev) and S-phase (p29of40CtDev) of the cell cycle. In the previous 7 pattern analysis (p7CtDev), the RG and cycling cell pattern p5of7CtDev showed enrichment for genes involved in monogenic microcephaly ( $p=4.1\text{e-}9$ ), but only a modestly



significant enrichment for genome-wide association with total cortical surface area ( $p=0.001$ ). The G2M pattern p7of40CtxDev is present in a much more limited subpopulation of progenitors than p5of7CtxDev, yet showed an even greater enrichment for both monogenic microcephaly ( $p=4.9e-16$ ) and total cortical surface area ( $p=3.8e-6$ ). In the Micali 2023 macaque data this dissection of the cell cycle can be seen independently in radial glia (lower horizontal loop with distinct areas of high p7 and high p29) and intermediate progenitors (upper horizontal loop with distinct areas of high p7 and high p29). These results indicate that by elevating the resolution of transcriptome decomposition, it is possible to identify increasingly specific transcriptomic elements within cell subtypes and their association with brain structure and disease. Additionally, gene loadings for this oRG signature were enriched in genes involved in cholesterol homeostasis ( $p=6.2e-12$ ). Along with the canonical oRG markers, KLF6, a gene recently implicated in evolutionarily new mechanisms in cholesterol metabolism specifically in human oRG cells (129) is also among the top ranked genes in p27of40CtxDev. Supplementary Table 3 contains the full set of gene loadings across all 40 patterns and Supplementary Table 4 contains the full set enrichments across all 40 patterns. In the lower panels of A, we plot individual single-cell scores from each of the highlighted jointNMF signatures (Y-axis) across pseudotime calculated with Monocle 3 (X-axis), colored by the consensus MetaMarker cell type calls. **B]** Boxplots of projected single-cell embeddings for p27of40CtxDev (oRG) separated by cell type calls made by original authors. For LaManno 2021 (11) and Micali 2023 (12) the original author cell type calls were used. Trevino 2021 (7) did not make detailed cell type calls (including no oRG calls). Therefore, we added some detail to these calls for both this boxplot and the CellOracle analysis in the main Figure 3 (See “Trevino et al 2021 Cell type reannotation and CellOracle analysis” section of Methods). Boxplot characteristics: center line=median; box limits=upper and lower quartiles; whiskers=1.5x interquartile range. **C] Left panels:** To further explore the individual genes involved in this partially conserved transcriptomic signature, we plotted each gene’s loading in p27of40CtxDev against the average expression of that gene in the mouse, macaque or human cells that have highest levels of this oRG signature (Methods). In each of the 3 species, we observed the expected positive correlation between the p27 loadings and expression levels in cells high in p27. Consistent with a cell type of evolutionarily increasing cohesiveness, this correlation grows from mouse to macaque to human ( $p<2.2e-16$  in each case), with the relative expression of canonical markers of human oRG cells increasing along this same evolutionary trajectory (blue labels). In the mouse specifically, there are many genes that have near-zero loadings in p27of40CtxDev (i.e. genes NOT involved in the p27 transcriptomic signature of oRGs) while still having high expression in the cells with the highest p27 levels in the developing murine neocortex (red highlighted genes at low X-axis values). In contrast, in both the macaque and human oRG cells, no genes with low p27 loadings are expressed at high levels. This raises the possibility that these genes are part of a transcriptomic program that has been shut down in oRG cells since the rodent-primate divergence. The transcriptional repressor FOXN3 is one of the genes that is low in p27 gene loadings and high in mouse but not in macaque or human cells (in green). Strikingly, FOXN3 target genes are enriched among high p27 gene loadings (targets from MsigDB;  $p=3.9e-20$ , Table S2). The gray curve is a loess fit of the average expression of genes across the magnitude of gene loadings in this pattern. **C] Right panels:** To explore this further, we integrated scRNA-seq and scATAC-seq data from the Trevino human dataset (7) in CellOracle (108) where it is possible to estimate regulatory networks and test the simulated effects of TF perturbations on cell identity *in silico*. We found that FOXN3 knock-out (KO) simulation produced many cell state transitions to the oRG state, while over-expression (OE) lead to its near disappearance. This suggests the possibility that high FOXN3 expression in mouse progenitors destabilizes the oRG state, consistent with their sparsity in the rodent brain. This data and the role of FOXN3 as a transcriptional repressor and proliferation inhibitor (130, 131, 132) are consistent with a model in which primate oRG cells, where FOXN3 expression is low, allow the increased expression of FOXN3 targets, resulting in the observed enrichment of these targets in p27 gene weights, and the stabilization of the proliferative oRG cell state in primate and human neocortex. Hence, it appears that the evolution of the oRG transcriptomic program in primates may have involved both subtraction from and addition to the ancestral neural progenitor transcriptomic program. Custom cell type calls in the Trevino 2021 data (see Methods): vRG=ventricular radial glia, oRG=outer radial glia, tRG=truncated radial glia, mGPC=multipotent glial precursor cells, Astro=astrocytes, OPC=oligodendrocyte progenitor cells. **D]** Comparison of projected expression patterns of the RG/cycling p5of7CtxDev (from Figure 1 & 2) and p27of40CtxDev (from the higher resolution decomposition from Figure 3) neural progenitor transcriptome signatures. To explore the expression of these signatures more widely (especially in gliogenesis), the La Manno 2021 and Trevino 2021 data in this panel D contain many more cells than were used in the original decomposition (Figures 1 & 3A). Black arrows indicate expression in neurogenic precursors, green arrows indicate expression in gliogenic precursors and glia. While p5 is expressed in both neurogenic and gliogenic precursors in mouse (10), macaque (12), and human (133) cortex, p27 has high expression levels in neurogenic precursors only in primate (Figure 3D) and human cortex (Figure 3D and S3D). Explore more at [NeMOLinkS04](#) by selecting p27. Plots in panels A, D and E are images downloaded directly from NeMO Analytics. **E]** Projection of microarray data from laser capture microdissected macaque (36) and human (37) developing brain tissue into the oRG signature p27of40CtxDev. This highlights the early expression of p27 in the VZ that was also noted by Pollen et al. 2015 (30). **F]** All 40 individual patterns in the p40CtxDev jointNMF decomposition. Individual gene weights underlying these patterns and their gene set enrichments are in Table S3 and all 40 patterns can be explored across mammalian neocortical developmental data at [NeMOLink08](#). Images in panel A, D and E were created from NeMO Analytics image downloads.

Supplemental Figure 4 - related to main Figure 4





C

Jorstad 2023\*  
Hs, Adult, multi-  
region

Bakken 2021\*  
Hs, Adult, Motor  
Cortex

Bakken 2021  
Marmoset, Adult,  
Motor Cortex

Bakken 2021  
Mouse, Adult, Motor  
Cortex

Ma 2022  
Human, Adult, DLPFC

Ma 2022  
Chimp, Adult, DLPFC

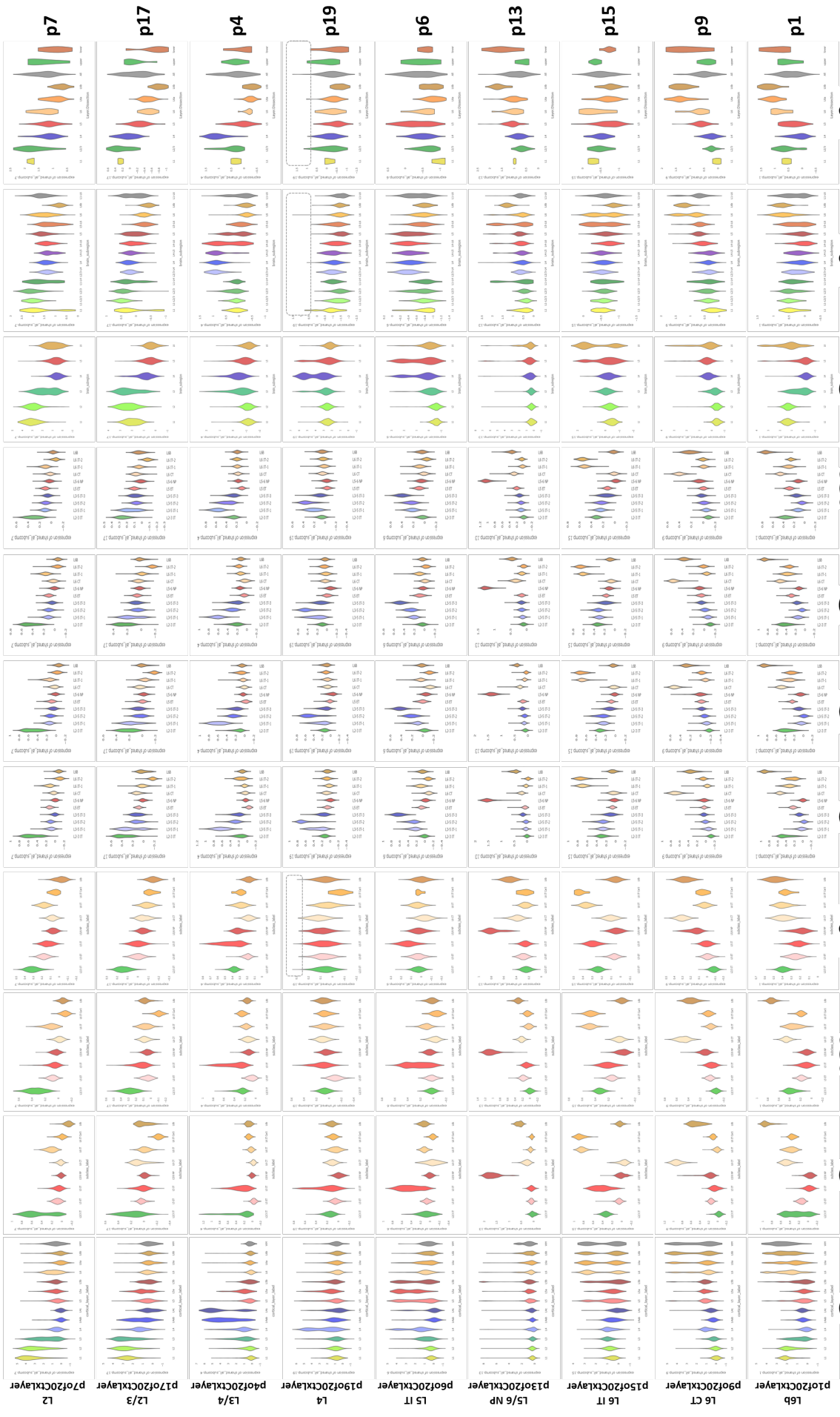
Ma 2022  
Macaque, Adult,  
DLPFC

Ma 2022  
Marmoset, Adult,  
DLPFC

Hodge 2019  
Human, Adult, MTG

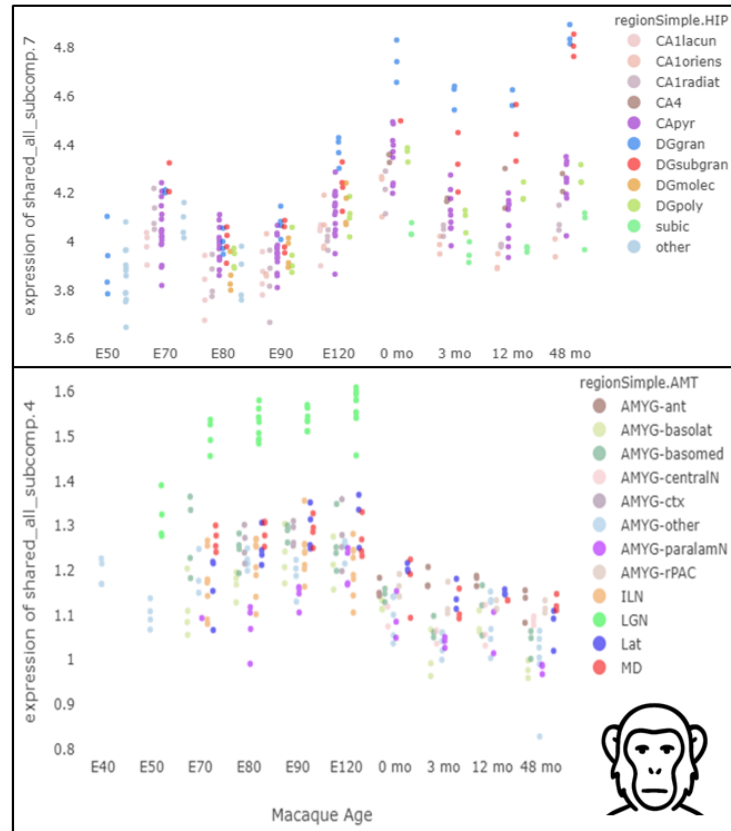
Tasic, 2018  
Mouse, Adult, Visual  
& motor Cortex

Tasic, 2016  
Mouse, Adult, Visual  
cortex



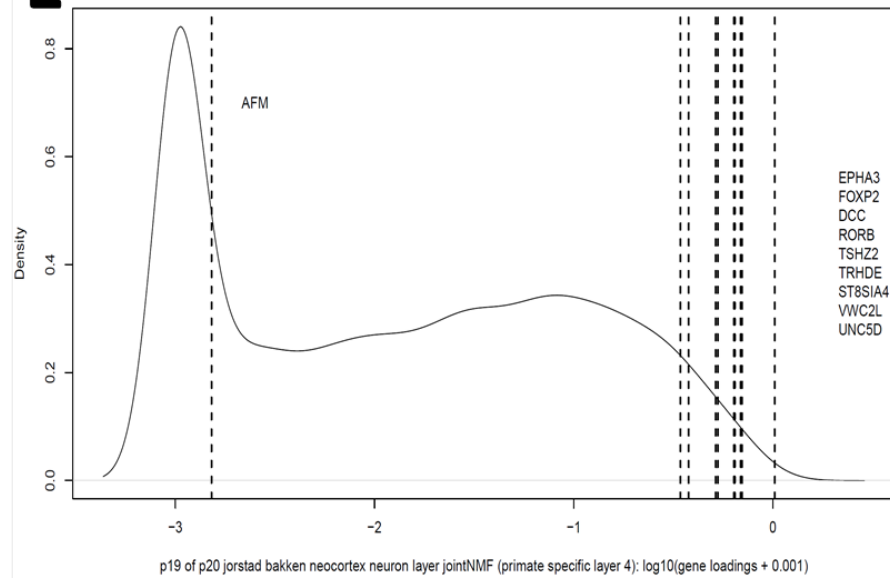
D

Bakken 2016  
MacMul; LMD+microarray



E

Chen et al Primate specific layer 4 genes in the distribtion of p19 gene loadings - geneSetTest:  $p=4.5e-06$





Jorstad 2023, snRNA-seq  
Hs, Adult

Bakken 2021, snRNA-seq  
Hs, Adult, Motor Cortex

L6b  
p1of20CtxLayer



L3/4  
p4of20CtxLayer



L5 IT  
p6of20CtxLayer



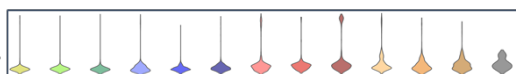
L2  
p7of20CtxLayer



L6 CT  
p9of20CtxLayer



L5/6 NP  
p13of20CtxLayer



L6 IT  
p15of20CtxLayer



L2/3  
p17of20CtxLayer



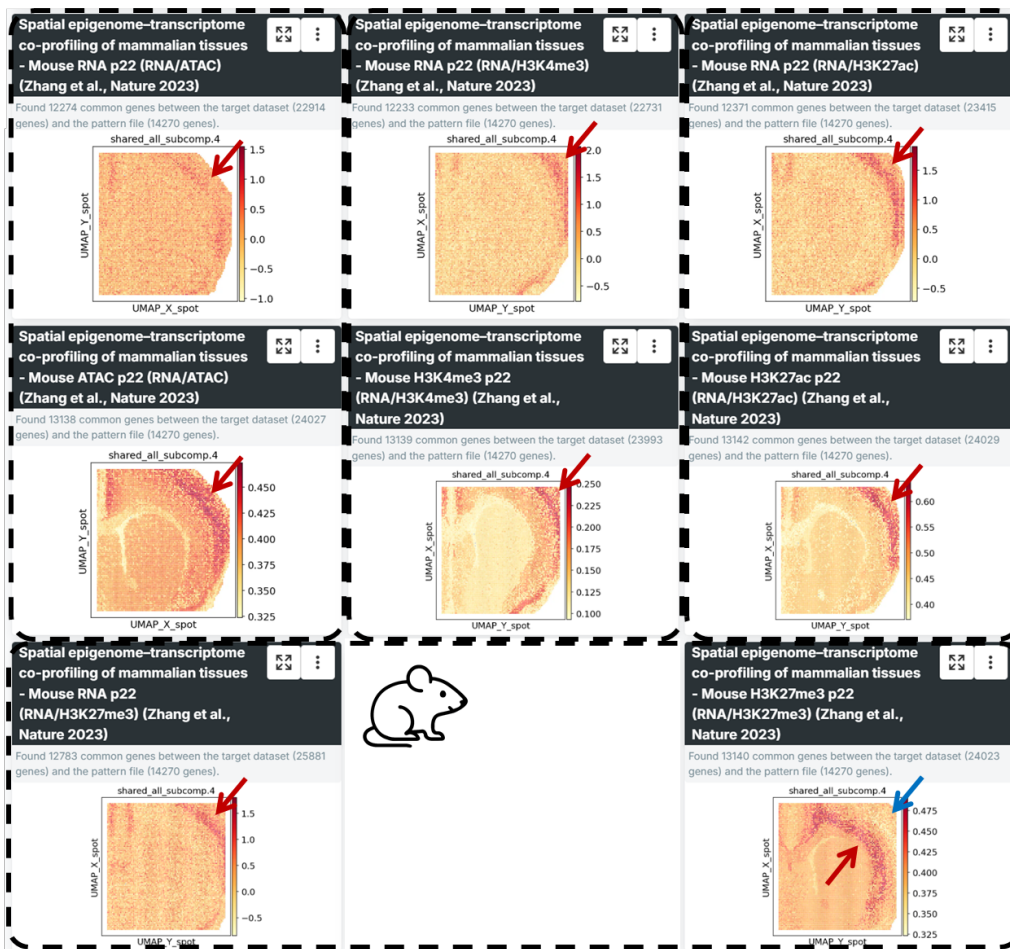
L4  
p19of20CtxLayer



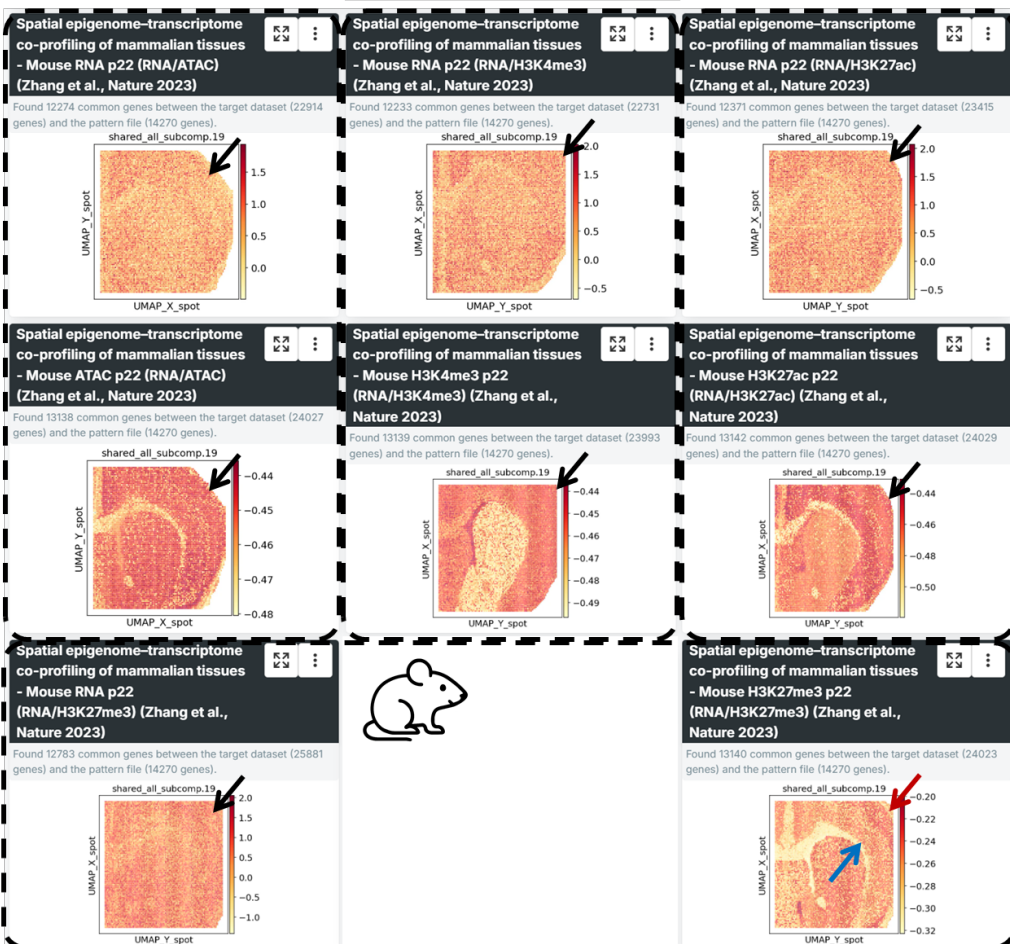
L1 L2 L3 L4 L4ab L4c L5 L5a L5b L6 L6a L6b WM

L2/3 IT L5 ET L5 IT L5/6 NP L6 CT L6 IT L6 IT Car3

G  
p4



H  
p19





**Supplemental Figure 4 - related to main Figure 4: A]** Violin plots of cell embeddings for the two adult human snRNA-seq (SMART-seq) datasets used in the jointNMF() decomposition in Figure 4 (Jorstad 2023, 38 & Bakken 2021, 39). Embeddings for 9 of the 20 total transcriptomic signatures (p20Layers) are shown. These are the same 9 focused on in main Figures 4 & 5, where projections in other datasets are shown. Here embeddings from individual donors are shown separately to show consistency across samples. **B]** Violin plots of cell embeddings from the same Jorstad 2023 (38) data as in panel A, here rearranged to show embeddings from individual regions to show consistency across neocortical areas. **C]** Violin plots for a wide collection of snRNA-seq datasets demonstrating the species and laminar specificity of the transcriptomic signatures from adult human cortex focused on in Figure 4 & 5. Dashed boxes indicate where the primate and human specific layer 4 pattern, p19of20CtxLayer, is absent in the mouse neocortex. Asterisks (\*) indicate the two datasets used in the original decomposition. **D]** Projection of LMD coupled microarray expression data from non-cortical regions of the developing macaque brain into p7 and p4of20CtxLayer which showed shared signatures with hippocampal and thalamic neurons, respectively, in mouse spatial transcriptomic data in main Figure 4. The LMD data confirm that these transcriptomic signatures are also shared across these same brain regions in macaque, with p7 at high levels specifically in the dentate gyrus granule layer, and p4 high in the lateral geniculate nucleus of the thalamus. These projections clearly demonstrate that multiple human adult laminar neuron transcriptome patterns in the neocortex are shared with neurons in diverse regions of the brain and these shared programs are conserved across mammals, consistent with recent scRNA-seq mapping of neuronal types across the cortex and hippocampus in the mouse brain (134) where numerous neuronal types were observed to be common across these two brain structures. **E]** Ma 2022 (40) recently observed primate specific expression of FOXP2 (a gene that has been implicated in human language development and neuropsychiatric disease; 135, 136) in excitatory neurons of layer 4. FOXP2 is ranked in the top 1% of genome-wide loadings for the primate-specific layer 4 pattern p19. Similarly, Chen 2023 (41) reported primate-specific layer 4 neuronal cell types in a recent study of spatial transcriptomics in the adult macaque brain. Interrogation of the genes reported in the Chen layer 4 signal in p19 gene loadings revealed a significant enrichment of high values in this small group of 11 genes ( $p=4.5e-6$ ), which again included FOXP2. Hence, the primate-specific layer 4 pattern, p19, that we have defined here is likely detecting the same FOXP2-related layer 4 signal as that described in Chen 2023 (41) and Ma 2022 (40). Panel E depicts a density plot of gene loadings in the p19of20CtxLayer transcriptomic signature of primate and human specific layer 4 neurons. The loadings in for the 11 genes reported by Chen 2023 (41) to be specifically in layer 4 neurons of primates and humans are shown as vertical dashed lines and show significant enrichment among high p19 loadings in a Wilcoxon rank sum test (implemented via the geneSetTest() function in the limma package in the R statistical language;  $p=4.5e-6$ ). **F]** The complete set of 20 transcriptomic patterns in the p20CtxLayer decomposition, shown as cell embeddings aggregated across donors and regions in the Jorstad 2023 (38) and Bakken 2021 (39) data that were used in the decomposition. Original author cell type calls were used. **G&H]** Screen shots of from NeMO Analytics depicting projection of multi-modal spatial -omics data from mouse (42) into human neuronal layer 4 maturity signatures: **G]** p4 (early maturing, mammalian conserved layer 4) and **H]** p19 (primate-specific layer 4). As expected for the conserved signature, p4, RNA expression and active histone marks show a clear signal in layer 4 of the mouse, while the repressive mark, H3K27me3, shows low levels in layer 4 and high levels in other deeper layers. For the primate-specific, p19, RNA expression and active histone marks show no specific signal, while the repressive mark, H3K27me3, is depleted from lower layers and specifically elevated in layer 4, indicating the active repression of this primate-specific signature in the mouse neocortex. Dotted boxes show coupled data modalities studied in the same tissue section. Black arrows indicate absence of specific signal, red arrows indicate high signals, and blue arrows indicate depleted signals. These results can be viewed live on NeMO Analytics at: [NeMOLinkS05](#).

**Additional notes on the regulation of layer-specific neuronal maturity signatures in Figure 4:** RORB and CUX2 are highly ranked in the p4 gene weights (2nd and 9th, respectively; Table S3). RORB is a canonical layer 4 marker and CUX2 is an upper layer marker, making this TF combination consistent with the precise localization of this signature to superficial layer 4, (spatial transcriptomics in Figure 4B). Interestingly, RORB (but not CUX2) is also highly ranked the p19 signature (347th), which peaks just below p4 in layer 4 proper, while CUX2 is also highly ranked in p17 (18th) and p7 (12th) that both show highest levels in layer 2/3.

As we point out in the main text and Figure S4E above, previous reports corroborate the observation of p19 as a primate-specific layer 4 signature (40, 41). The L3-5 IT-2 neuronal cell type identified in Ma et al. is particularly high in p19, as can be seen here via projection of our layer specific transcriptomic signatures in the adult neocortex data collection in NeMO Analytics ([NeMOLinkS06](#)). Using CellOracle-derived cell type specific gene regulatory networks in the adult human dorsolateral prefrontal cortex from Ma et al., we found that the FOXP2, RUNX2, and MEF2A regulons in L3-5 IT neurons are enriched in the primate-specific p19 (again implicating FOXP2 in this primate-specific signature, and highlighting complex multi-TF regulation), while MEF2A and RORA regulons in L3-5 IT neurons are linked to the conserved p4 signature. We also examined FOXP2 target genes from two overexpression studies (137, 138) by projecting them into the adult neocortex datasets ([NeMOLinkS07](#)). Although these FOXP2 targets show some layer enrichment, they exhibit less specificity compared to the more sharply defined layer- and species-specific patterns captured by our joint decomposition (as we also saw in the projection of FOXP2 regulons in L3-5 IT neurons from (40) - [link](#)), suggesting FOXP2 is one of multiple regulators of p19, as well as additional cortical layer neuronal maturity signatures.

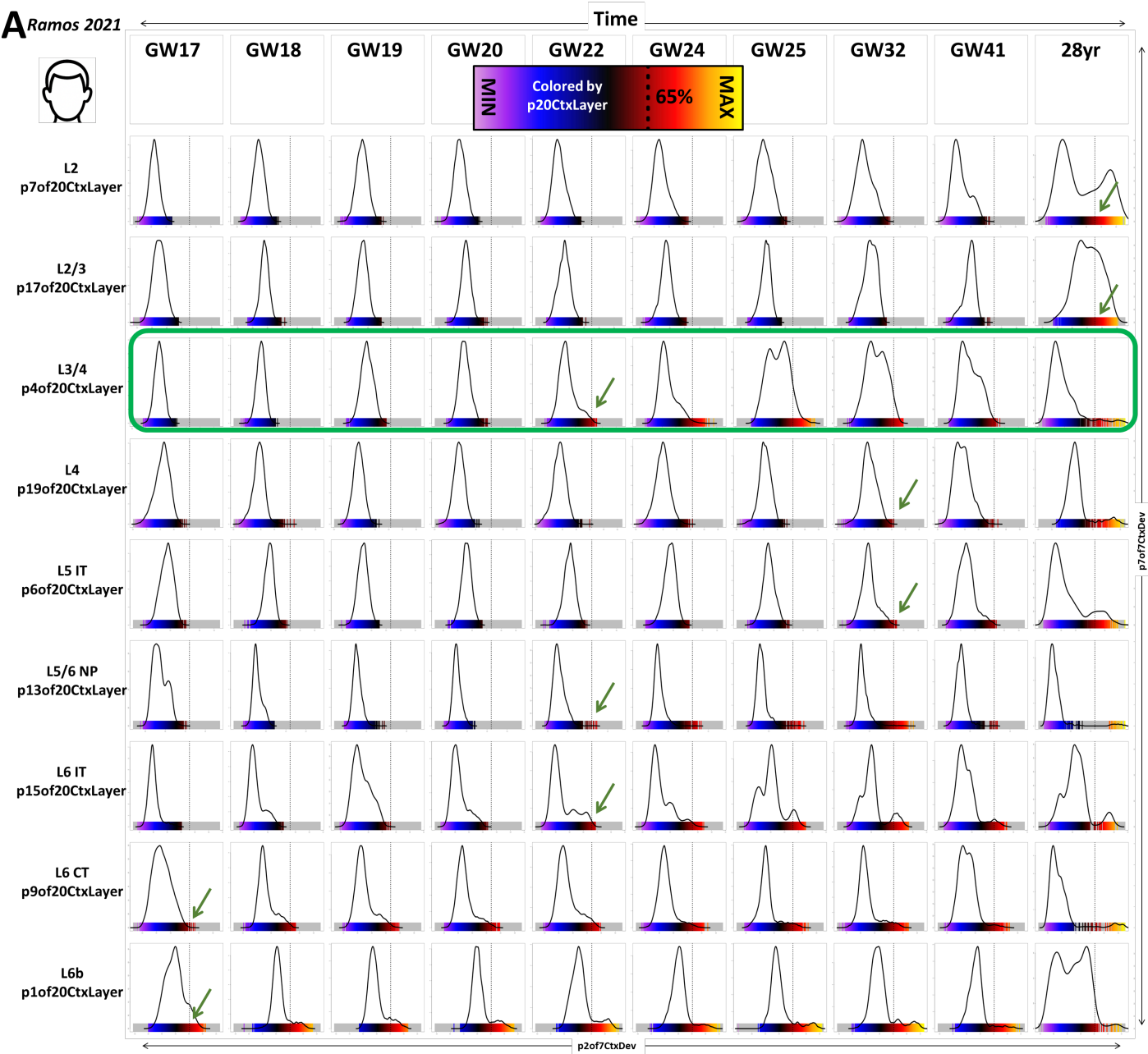
The enrichment of MEF2A regulons in both layer 4 patterns (p4 and p19), encouraged us to further examine MEF2A in this context. In addition to the enrichment of L3-5 IT neuronal MEFA2 regulons in p4 and p19, several MEFA2 regulons showed broad enrichment across our neuronal maturity patterns. To explore this from another perspective, we captured a MEFA2 gene expression module (139) shown to have delayed expression in human cortex when compared to macaque and chimpanzee trajectories. Genes in this synaptic activity related gene expression module were enriched in all our neuronal maturity signatures, indicating that, like FOXP2, MEFA2 plays a role in many neuronal maturity programs. Projection of this MEFA2 module across many datasets can be seen ([NeMOLinkS7b](#)) reaching its maximum level at birth in the macaque and at approximately 2 years in the human. Clearly, much more work is required to understand the GRNs underlying these neuronal maturity programs – we hope the observations we have made here, along with the individual gene level quantification for each signature produced by the joint decomposition, can facilitate further exploration by other groups.



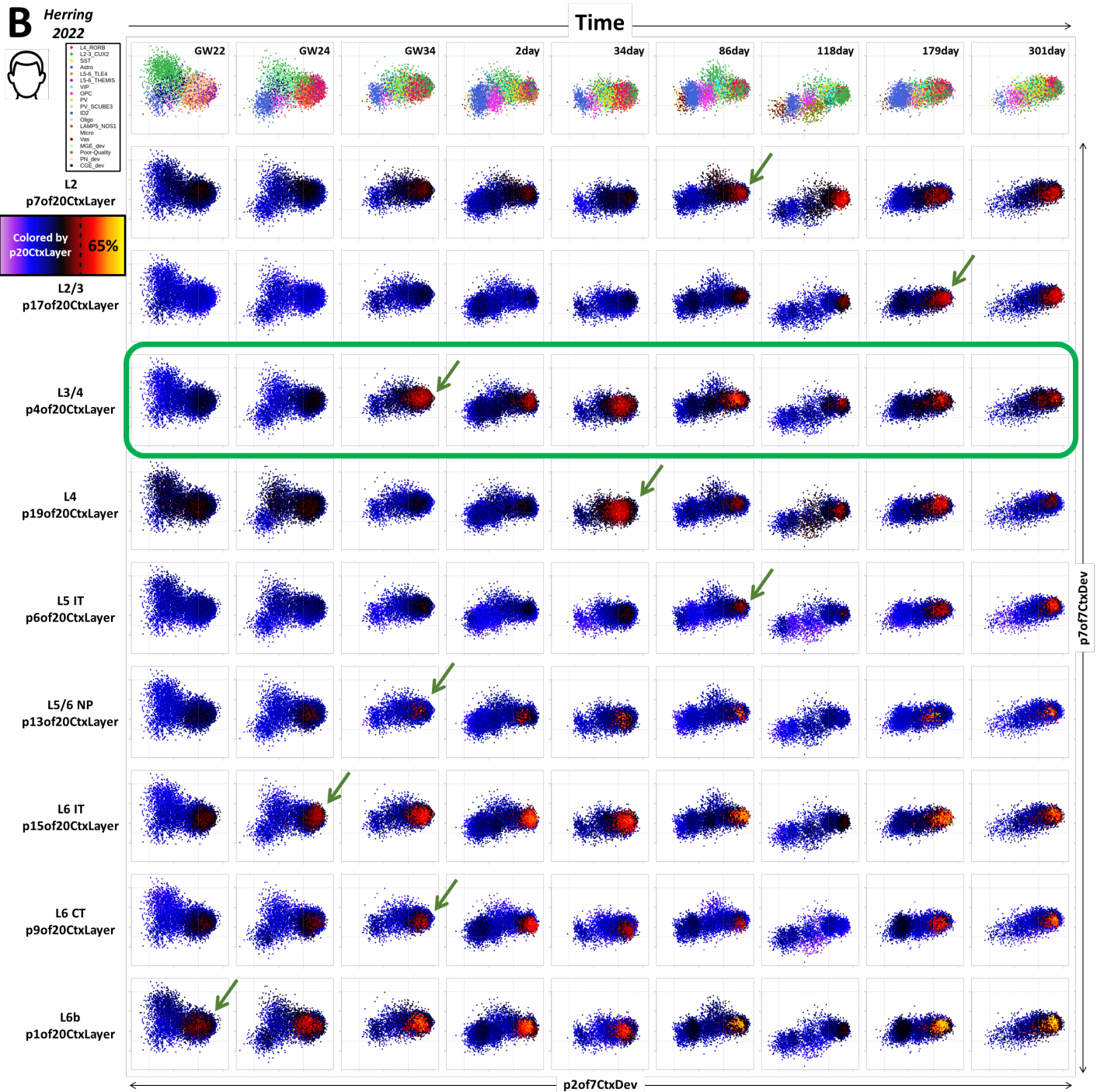
**Supplemental Figure 5 - related to main Figure 5: A]** Projection of macaque LMD-coupled microarray data into the p20CtxLayer patterns. The laminar specificity and protracted timescales of these maturational trajectories in the non-human primate are made clear in the laser capture microdissection (LMD)-coupled expression analysis spanning pre- and post-natal macaque neocortical development from Bakken 2016 (36). Additional observations in this unique dataset include the linking of fetal subplate and mature layer 6b molecular identities in p1of20CtxLayer (consistent with 140), V1-specific enrichment of the p9 layer 6, and p4 layer 3/4 identities, all of which play out over years of postnatal development (arrows). These transcriptomic signatures can also be explored online (including individual sample annotation in the LMD data) at: [NeMOLinkS05](#). **B]** Line plots of the mean projected single-cell embeddings in neurons across pre- and post-natal life for each adult layer-specific neuronal pattern (9 of 20 patterns in p20CtxLayer) in the Herring 2022 (43) snRNA-seq data split into 4 neuron subtypes identified by the original authors. Solid lines join means at each age/donor. Dashed lines are local fits to the individual age means made with the loess() function in the R statistical language. **C]** Originally published UMAP of fetal neurons colored by projected single-cell embeddings for each adult layer-specific neuronal pattern (9 of 20 patterns in p20CtxLayer) in the Herring 2022 (43) snRNA-seq data. The adult laminar-specific neuronal transcriptomic signatures (especially the deep-layer signatures) have strikingly clear specificity in particular cell populations at these early stages of development. For example, the 3 patterns which show high levels in TLE4+ neurons (p1, p9, and p13of20CtxLayer), all show high levels in clearly distinct sub-populations of the neurons classified as TLE4+ by Herring 2022 (arrows). While following the general pattern of building laminar-specific expression over time within a particular neuronal type, two of the three patterns in TLE4+ neurons, p1of20CtxLayer (L6b) & p13of20CtxLayer (L5/6 NP), display particularly unique maturational trajectories (see also main Figure 5B). In fetal and early postnatal development, these patterns each rise specifically within distinct sub-populations of TLE4+ neurons and fall in other TLE4+ neurons. Interestingly, levels of p13 fall across development and throughout adulthood in all neuron types with the exception of the specific TLE4+ L5 NP subtype in which they are finally enriched (see also main Figure 5B). Hence, while all layer identities arise from increasing expression levels in a specific neuronal cell type, some also emerge via repression in other neuronal types. **D]** The Herring 2022 (43) dataset captured very few nascent neurons (high in p7of7CtxDev & low in p2of7CtxDev) and many maturing neurons (low in p7 and high in p2). This is perhaps due to tissue dissection bias or other factors that we have not determined. Comparing panel C to plots in panel D, it appears that the adult layer-specific neuronal transcriptomic signatures begin to emerge in fetal neurons that have shut down p7 and induced p2, as earlier analysis of patterns p7of7CtxDev & p2of7CtxDev in Figure 2 & S2 indicated. This is explored in much greater detail in Figure 6 & S6. With the exception of the line plots and additional labels, this entire figure was created from NeMO Analytics image downloads. Original author cell type calls were used. Exploring even earlier data from human neocortical development in Braun et al. (19) and additional studies, indicates that in the first trimester, there is virtually no sign of the emergence of the neuronal maturity signatures that we defined from adult snRNA-seq data ([NeMOLinkS08](#)). **E]** Additional layer defining TFs in the Herring 2022 (43) dataset.



Supplemental Figure 6 - related to main Figure 6

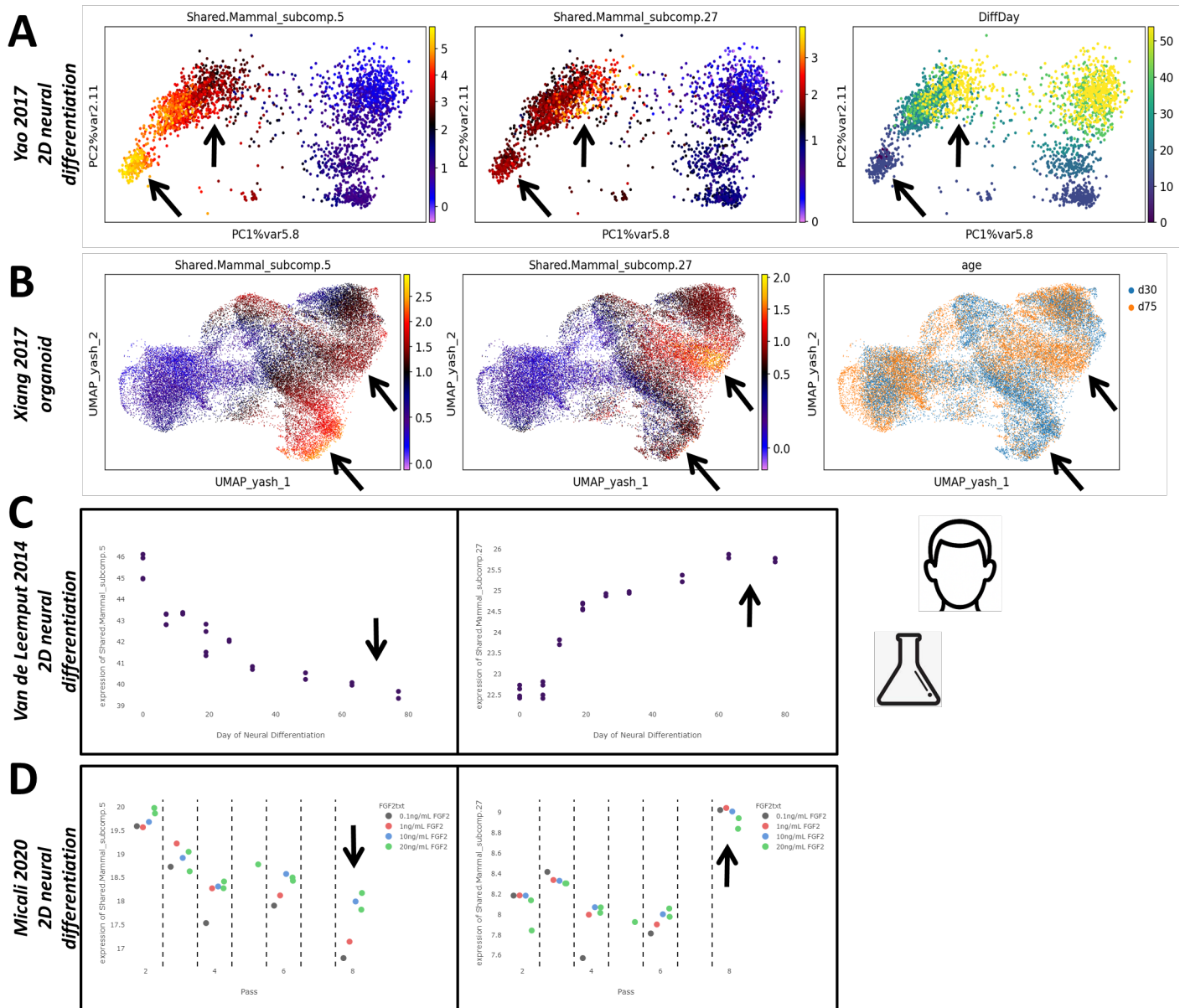


**Supplemental Figure 6 - related to main Figure 6: A]** This collection of plots is constructed just as the matrix of plots in Figure 6, with each mature neuronal pattern in a row and developmental ages across the columns, but here focusing on only the emergence of the adult transcriptomic signatures. The density plots show the distribution of projected single-cell embeddings for each of the adult neuronal transcriptional signatures at each age in fetal development and a single adult time point in the Ramos 2021 (8) data. Each individual single-cell embedding is plotted as a vertical color strip beneath the density plot. Green arrows indicate the age at which cells first begin to surpass 65% of the maximal level for each signature. In the color scale here (same as in Figure 6) this is where the coloring of samples begins to turn visibly red. As observed in Figure 6, the conserved layer 4 pattern, p4of20CtxLayer, appears earlier than other cells at similar laminar positions (green box). As in Figure 6, this includes only cells of the excitatory neurogenic lineage. To remove low quality cells, in both Figure 6 and here in Figure S6A, we performed an additional QC step, removing cells with fewer than 1250 gene detected or greater than 5% mitochondrial reads.



**Supplemental Figure 6 - related to main Figure 6: B]** Recreation of the same plots of projected p7 vs. p2of7CtDev embeddings colored by projected p20CtLayer embeddings in Figure 6 using data from Ramos 2021 (8), here using data from Herring 2022 (43). Because Herring 2022 (43) did not make any progenitor cell type calls, we included all cells to ensure we did not miss any of the excitatory neurogenic trajectory. As in the Ramos 2021 (8) data in Figure 6, here the conserved layer 4 pattern, p4of20CtLayer, appears earlier than other cells at similar laminar positions (green box). The same 65% threshold that was applied in the Ramos data in Figures 6 and S6A is applied here in the Herring data, however, the Ramos data included many more early fetal time points. This difference in age ranges covered likely inflates the estimated of ages of signature emergence in the Herring 2022 (43) data. Original author cell type calls were used. NOTE on earlier time points: We have examined the projection of these adult human layer-specific neuronal maturation signals in 1st trimester human data (19) assess the emergence of these neuronal maturity signals in early development. Consistent with the timing of emergence we see in Figures 6 and S6, in the first trimester data there is no sign of the emergence of the neuronal maturity signatures defined from adult snRNA-seq data. The neuronal maturity patterns can be explored in the 1<sup>st</sup> trimester data (among many other datasets) at [NeMOLinkS08](#).

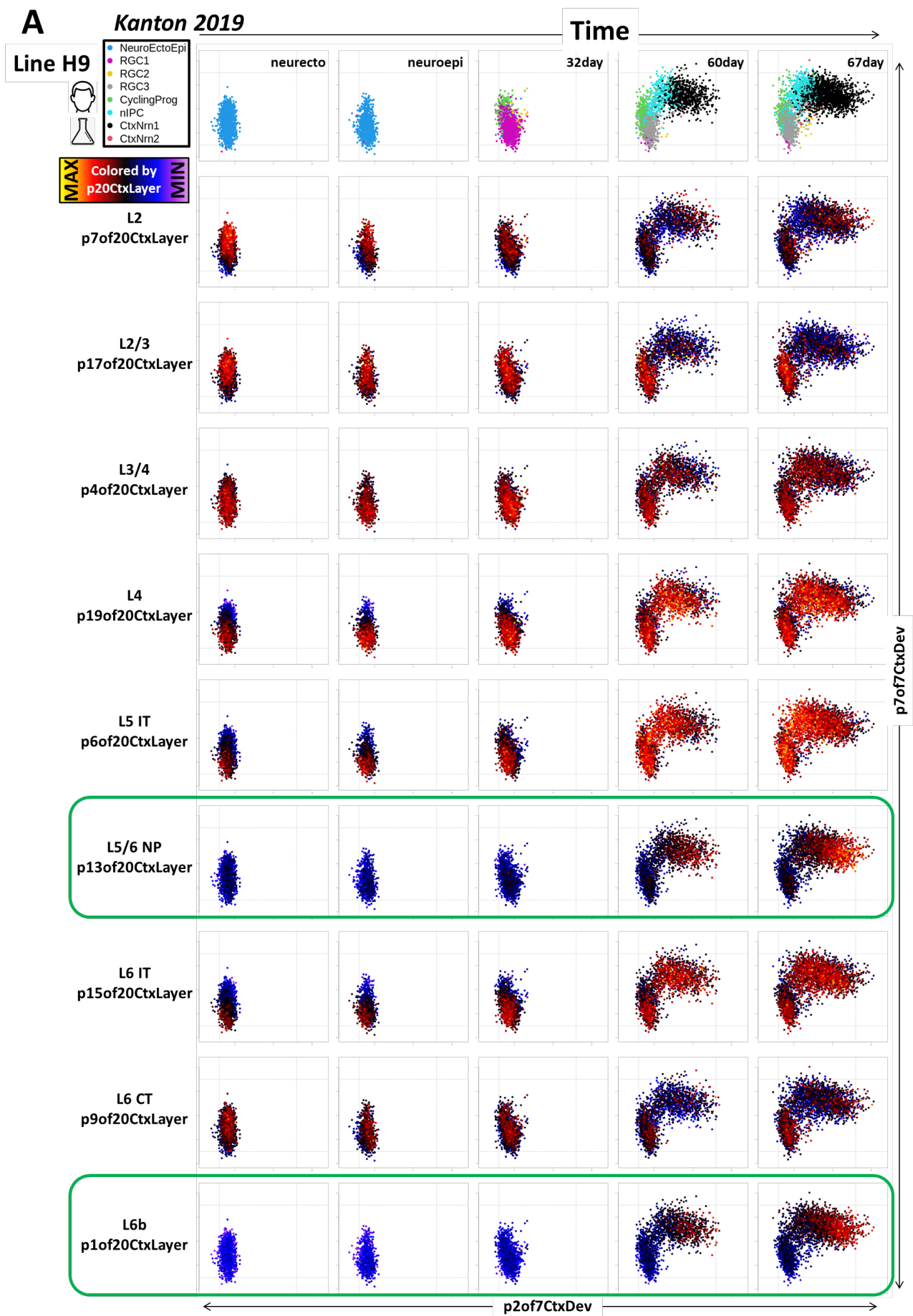
## Supplemental Figure 7 - related to main Figure 7



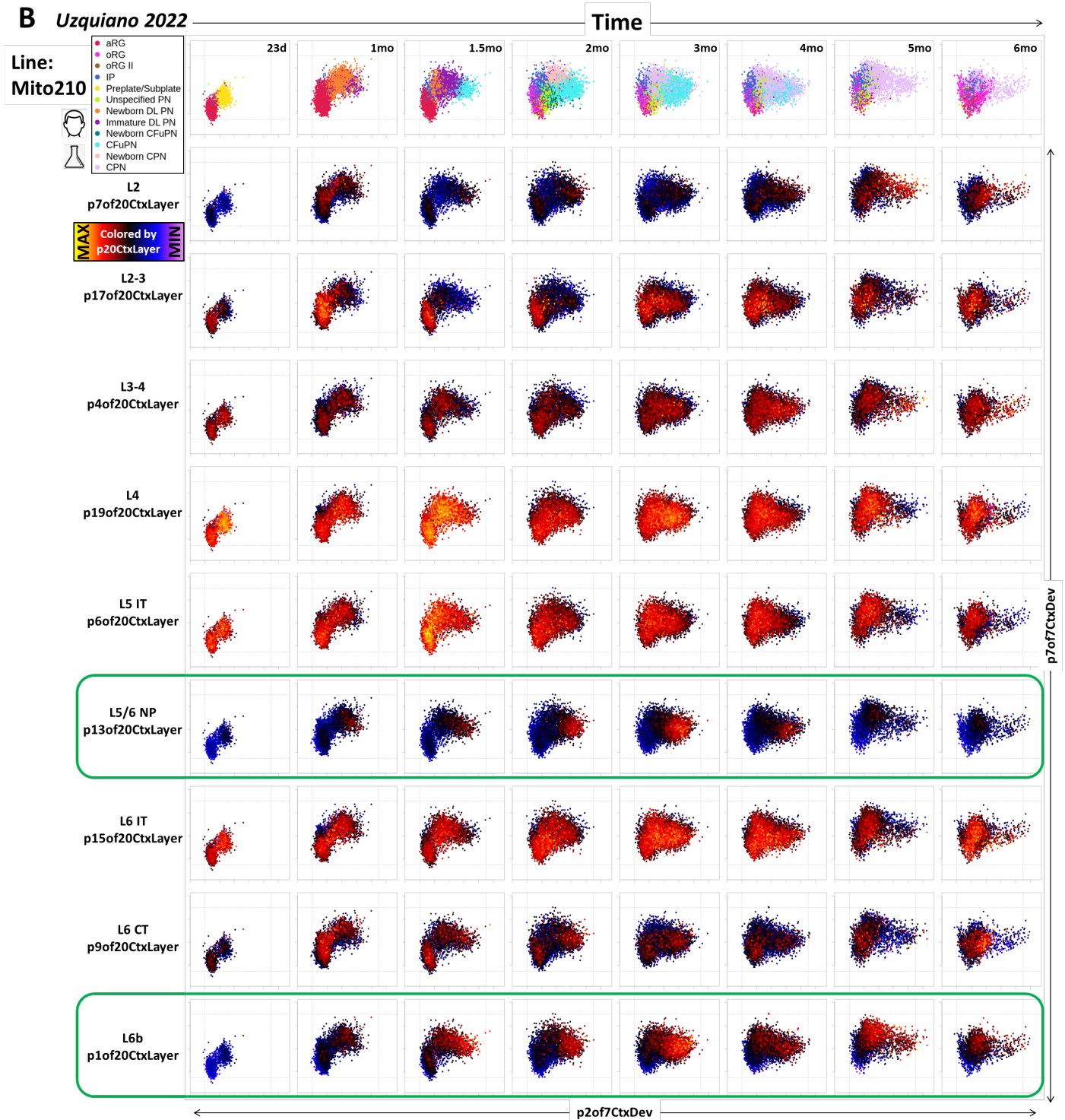
**Supplemental Figure 7 - related to main Figure 7: p5of7CtxDev and p27of40CtxDev differ systematically in their temporal emergence *in vitro*.** As noted using *in vivo* data in Figures 1, 2, and S3, p5of7CtxDev is expressed at particularly high levels in early dividing RG, while p27of40CtxDev is a transcriptomic signature of later oRG and also high in gliogenic precursors that emerge following neurogenesis. Here we show with projection of data from iPSC-derived neural differentiation systems, that these two transcriptional programs emerge in a similar succession *in vitro*: **A]** Yao 2017 (141) shows p5of7CtxDev highest in neural precursor cells growing in 2-dimensional culture at day 7, and p27of40CtxDev is highest in a small population of precursor cells at day 54. **B]** Xiang 2017 (61) shows high expression of p5of7CtxDev in precursors at day 35 of organoid culture, while p27of40CtxDev is highest in a distinct population of precursors, later at day 75. **C]** Van de Leemput 2014 (142) using bulk RNA-seq showed a similar temporal trend, with p5of7CtxDev high early and p27of40CtxDevo high later. **D]** Micali 2020 (143) using bulk RNA-seq across the passaging of human neural cells shows highest levels of p5of7CtxDev in early passages and highest p27of40CtxDev at later passages. Arrows indicate time points at which the expression of p5of7CtxDev and p27of40CtxDev differ. All panels are images downloads from NeMO Analytics.



Supplemental Figure 8 - related to main Figure 8







**Supplemental Figure 8 - related to main Figure 8: A]** Main Figure 8 shows the organoid time course for 1 human iPSC line (409b2) from the Kanton 2019 study (48), here we show the time course for the H9 human ESC line. The same two deep layer transcriptomic signatures appear systematically in the H9 line here as appeared in the 409b2 line (green boxes). **B]** As in both hPSC lines from the Kanton 2019 data, the same two deep layer transcriptomic signatures appear systematically here in the Uzquiano 2022 (15) organoid time course (green boxes). The Uzquiano 2022 time course is much longer than that in Kanton 2019. The L5/6 NP pattern (p13of20CtxLayer) appears to peak at 3 months, decreasing thereafter and almost disappearing by 6 months. This is likely due to the rise of the gliogenic lineages along with possible necrosis of portions of the organoid that were dedicated to specific neuronal populations. Notably, the L6b pattern (p1of20CtxLayer) also appears to peak at 3 months, but shows a less dramatic decrease thereafter. Original author cell type calls were used: NeuroEctoEpi=neurectodermal and neuroepithelial states, RGC=radial glial cells, CyclingPrg=cycling neural progenitors, nIPC=neuronal intermediate progenitor cell, CtxNrn=cortical neuron, aRG=apical radial glia, oRG=outer radial glia, IP=intermediate progenitor, DL=deep layer, PN=projection neuron, Cfu=corticofugal, CPN=cortical projection neuron.

## Supplemental References

112. Smeland, O.B., Frei, O., Fan, C.C., Shadrin, A., Dale, A.M., Andreassen, O.A. The emerging pattern of shared polygenic architecture of psychiatric disorders, conceptual and methodological challenges. *Psychiatric genetics*. 29, 152-159 (2019). PMID: 31464996
113. Owen, M.J., Legge, S.E., Rees, E., Walters, J.T.R., O'Donovan, M.C. Genomic findings in schizophrenia and their implications. *Molecular psychiatry*. 28, 3638-3647 (2023). PMID: 37853064
114. Gordovez, F.J.A, McMahon, F.J.. The genetics of bipolar disorder. *Molecular psychiatry*. 25, 544-559 (2020). PMID: 31907381
115. Jayaraman, D., Bae, B.I., Walsh, C.A. The Genetics of Primary Microcephaly. *Annual review of genomics and human genetics*. 19, 177-200 (2018). PMID: 29799801
116. Jin, S.C. Exome sequencing implicates genetic disruption of prenatal neuro-gliogenesis in sporadic congenital hydrocephalus. *Nature medicine*. 26, 1754-1765 (2020). PMID: 33077954
117. Jamuar, S.S., Walsh, C.A. Genomic variants and variations in malformations of cortical development. *Pediatric clinics of North America*. 62, 571-585 (2015). PMID: 26022163
118. Romero, D.M., Bahi-Buisson, N., Francis, F. Genetics and mechanisms leading to human cortical malformations. *Seminars in cell & developmental biology*. 76, 33-75 (2018). PMID: 28951247
119. Subramanian, L., Calcagnotto, M.E., Paredes, M.F. Cortical Malformations, Lessons in Human Brain Development. *Frontiers in cellular neuroscience*. 13, 576 (2019). PMID: 32038172
120. Coulter, M.E. et al. Regulation of human cerebral cortical development by EXOC7 and EXOC8, components of the exocyst complex, and roles in neural progenitor cell proliferation and survival. *Genetics in Medicine*. 22, 1040-1050 (2020). PMID: 32103185
121. Searles Quick, V.B., Wang, B., State, M.W. Leveraging large genomic datasets to illuminate the pathobiology of autism spectrum disorders. *Neuropsychopharmacology*, official publication of the American College of Neuropsychopharmacology. 46, 55-69 (2021). PMID: 32668441
122. Willsey, H.R., Willsey, A.J., Wang, B., State, M.W. Genomics, convergent neuroscience and progress in understanding autism spectrum disorder. *Nature reviews. Neuroscience*. 23, 323-341 (2022). PMID: 35440779
123. Clifton NE, Lin JQ, Holt CE, O'Donovan MC, Mill J. Enrichment of the Local Synaptic Translatome for Genetic Risk Associated With Schizophrenia and Autism Spectrum Disorder. *Biol Psychiatry*. 2024 May 1;95(9):888-895. doi: 10.1016/j.biopsych.2023.12.006. Epub 2023 Dec 14. PMID: 38103876.
124. Smeland OB, Bahrami S, Frei O, Shadrin A, O'Connell K, Savage J, Watanabe K, Krull F, Bettella F, Steen NE, Ueland T, Posthuma D, Djurovic S, Dale AM, Andreassen OA. Genome-wide analysis reveals extensive genetic overlap between schizophrenia, bipolar disorder, and intelligence. *Mol Psychiatry*. 2020 Apr;25(4):844-853. doi: 10.1038/s41380-018-0332-x. Epub 2019 Jan 4. Erratum in: *Mol Psychiatry*. 2020 Apr;25(4):914. doi: 10.1038/s41380-019-0456-7. PMID: 30610197; PMCID: PMC6609490
125. Luo W, Zhang C, Jiang YH, Brouwer CR. Systematic reconstruction of autism biology from massive genetic mutation profiles. *Sci Adv*. 2018 Apr 11;4(4):e1701799. doi: 10.1126/sciadv.1701799. PMID: 29651456; PMCID: PMC5895441.
126. Romero DM, Poirier K, Belvindrah R, Moutkine I, Houllier A, LeMoing AG, Petit F, Boland A, Collins SC, Soiza-Reilly M, Yalcin B, Chelly J, Deleuze JF, Bahi-Buisson N, Francis F. Novel role of the synaptic scaffold protein Dlgap4 in ventricular surface integrity and neuronal migration during cortical development. *Nat Commun*. 2022 May 18;13(1):2746. doi: 10.1038/s41467-022-30443-z. PMID: 35585091; PMCID: PMC9117333.
127. Birtele M, Del Dosso A, Xu T, Nguyen T, Wilkinson B, Hosseini N, Nguyen S, Urenda JP, Knight G, Rojas C, Flores I, Atamian A, Moore R, Sharma R, Pirrotte P, Ashton RS, Huang EJ, Rumbaugh G, Coba MP, Quadrato G. Non-synaptic function of the autism spectrum disorder-associated gene SYNGAP1 in cortical neurogenesis. *Nat Neurosci*. 2023 Dec;26(12):2090-2103. doi: 10.1038/s41593-023-01477-3. Epub 2023 Nov 9. PMID: 37946050.
128. Wen C, Margolis M, Dai R, Zhang P, Przytycki PF, Vo DD, Bhattacharya A, Matoba N, Tang M, Jiao C, Kim M, Tsai E, Hoh C, Aygün N, Walker RL, Chatzinakos C, Clarke D, Pratt H; PsychENCODE Consortium†; Peters MA, Gerstein M, Daskalakis NP, Weng Z, Jaffe AE, Kleinman JE, Hyde TM, Weinberger DR, Bray NJ, Sestan N, Geschwind DH, Roeder K, Gusev A, Pasaniuc B, Stein JL, Love MI, Pollard KS, Liu C, Gandal MJ; PsychENCODE Consortium.

Cross-ancestry atlas of gene, isoform, and splicing regulation in the developing human brain. *Science*. 2024 May 24;384(6698):eadh0829. doi: 10.1126/science.adh0829. Epub 2024 May 24. PMID: 38781368.

129. Moriano J, Leonardi O, Vitriolo A, Testa G, Boeckx C. A multi-layered integrative analysis reveals a cholesterol metabolic program in outer radial glia with implications for human brain evolution. *Development*. 2024 Aug 15;151(16):dev202390. doi: 10.1242/dev.202390. Epub 2024 Aug 27. PMID: 39114968; PMCID: PMC11385646.

130. Scott KL, Plon SE. Loss of Sin3/Rpd3 histone deacetylase restores the DNA damage response in checkpoint-deficient strains of *Saccharomyces cerevisiae*. *Mol Cell Biol*. 2003 Jul;23(13):4522-31. doi: 10.1128/MCB.23.13.4522-4531.2003. PMID: 12808094; PMCID: PMC164854.

131. Huot G, Vernier M, Bourdeau V, Doucet L, Saint-Germain E, Gaumont-Leclerc MF, Moro A, Ferbeyre G. CHES1/FOXN3 regulates cell proliferation by repressing PIM2 and protein biosynthesis. *Mol Biol Cell*. 2014 Mar;25(5):554-65. doi: 10.1091/mbc.E13-02-0110. Epub 2014 Jan 8. PMID: 24403608; PMCID: PMC3937083.

132. Sun J, Li H, Huo Q, Cui M, Ge C, Zhao F, Tian H, Chen T, Yao M, Li J. The transcription factor FOXN3 inhibits cell proliferation by downregulating E2F5 expression in hepatocellular carcinoma cells. *Oncotarget*. 2016 Jul 12;7(28):43534-43545. doi: 10.18632/oncotarget.9780. PMID: 27259277; PMCID: PMC5190042.

133. Nowakowski TJ, Bhaduri A, Pollen AA, Alvarado B, Mostajo-Radji MA, Di Lullo E, Haeussler M, Sandoval-Espinosa C, Liu SJ, Velmeshev D, Ounadjela JR, Shuga J, Wang X, Lim DA, West JA, Leyrat AA, Kent WJ, Kriegstein AR. Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science*. 2017 Dec 8;358(6368):1318-1323. doi: 10.1126/science.aap8809. PMID: 29217575; PMCID: PMC5991609.

134. Yao Z, van Velthoven CTJ, Nguyen TN, Goldy J, Sedeno-Cortes AE, Baftizadeh F, Bertagnolli D, Casper T, Chiang M, Crichton K, Ding SL, Fong O, Garren E, Glandon A, Gouwens NW, Gray J, Graybuck LT, Hawrylycz MJ, Hirschstein D, Kroll M, Lathia K, Lee C, Levi B, McMillen D, Mok S, Pham T, Ren Q, Rimorin C, Shapovalova N, Sulc J, Sunkin SM, Tieu M, Torkelson A, Tung H, Ward K, Dee N, Smith KA, Tasic B, Zeng H. A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell*. 2021 Jun 10;184(12):3222-3241.e26. doi: 10.1016/j.cell.2021.04.021. Epub 2021 May 17. PMID: 34004146; PMCID: PMC8195859.

135. Lai, C.S., Fisher, S.E., Hurst, J.A., Vargha-Khadem, F., Monaco A.P. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*. 413, 519-523 (2001). PMID: 11586359

136. Ferland, R.J., Cherry, T.J., Preware, P.O., Morrissey, E.E., Walsh, C.A. Characterization of Foxp2 and Foxp1 mRNA and protein in the developing and mature brain. *The Journal of comparative neurology*. 460, 266-279 (2003). PMID: 12687690

137. Konopka G, Bomar JM, Winden K, Coppola G, Jonsson ZO, Gao F, Peng S, Preuss TM, Wohlschlegel JA, Geschwind DH. Human-specific transcriptional regulation of CNS development genes by FOXP2. *Nature*. 2009 Nov 12;462(7270):213-7. doi: 10.1038/nature08549. PMID: 19907493; PMCID: PMC2778075.

138. Hickey SL, Berto S, Konopka G. Chromatin Decondensation by FOXP2 Promotes Human Neuron Maturation and Expression of Neurodevelopmental Disease Genes. *Cell Rep*. 2019 May 7;27(6):1699-1711.e9. doi: 10.1016/j.celrep.2019.04.044. PMID: 31067457; PMCID: PMC6794152.

139. Liu X, Somel M, Tang L, Yan Z, Jiang X, Guo S, Yuan Y, He L, Oleksiak A, Zhang Y, Li N, Hu Y, Chen W, Qiu Z, Pääbo S, Khaitovich P. Extension of cortical synaptic development distinguishes humans from chimpanzees and macaques. *Genome Res*. 2012 Apr;22(4):611-22. doi: 10.1101/gr.127324.111. Epub 2012 Feb 2. PMID: 22300767; PMCID: PMC3317144.

140. Hoerder-Suabedissen, A., Molnár, Z. Molecular diversity of early-born subplate neurons. *Cerebral cortex (New York, N.Y., 1991)*. 23, 1473-1483 (2013). PMID: 22628460

141. Yao Z, Mich JK, Ku S, Menon V, Krostag AR, Martinez RA, Furchtgott L, Mulholland H, Bort S, Fuqua MA, Gregor BW, Hodge RD, Jayabalu A, May RC, Melton S, Nelson AM, Ngo NK, Shapovalova NV, Shehata SI, Smith MW, Tait LJ, Thompson CL, Thomsen ER, Ye C, Glass IA, Kaykas A, Yao S, Phillips JW, Grimley JS, Levi BP, Wang Y, Ramanathan S. A Single-Cell Roadmap of Lineage Bifurcation in Human ESC Models of Embryonic Brain Development. *Cell Stem Cell*. 2017 Jan 5;20(1):120-134. doi: 10.1016/j.stem.2016.09.011. Epub 2016 Oct 27. PMID: 28094016; PMCID: PMC5261831.

142. van de Leemput J, Boles NC, Kiehl TR, Corneo B, Lederman P, Menon V, Lee C, Martinez RA, Levi BP, Thompson CL, Yao S, Kaykas A, Temple S, Fasano CA. CORTECON: a temporal transcriptome analysis of in vitro human cerebral cortex development from human embryonic stem cells. *Neuron*. 2014 Jul 2;83(1):51-68. doi: 10.1016/j.neuron.2014.05.013. PMID: 24991954.

143. Micali N, Kim SK, Diaz-Bustamante M, Stein-O'Brien G, Seo S, Shin JH, Rash BG, Ma S, Wang Y, Olivares NA, Arellano JI, Maynard KR, Fertig EJ, Cross AJ, Bürli RW, Brandon NJ, Weinberger DR, Chenoweth JG, Hoepfner DJ, Sestan N, Rakic P, Colantuoni C, McKay RD. Variation of Human Neural Stem Cells Generating Organizer States In Vitro before Committing to Cortical Excitatory or Inhibitory Neuronal Fates. *Cell Rep.* 2020 May 5;31(5):107599. doi: 10.1016/j.celrep.2020.107599. PMID: 32375049; PMCID: PMC7357345.