Cell type specific roles of FOXP1 during early neocortical murine development

Ana Ortiz, Fatma Ayhan, Matthew Harper, Genevieve Konopka*

Department of Neuroscience, Peter O'Donnell Jr. Brain Institute, UT Southwestern Medical Center, Dallas, TX, USA
*Correspondence: Genevieve.Konopka@utsouthwestern.edu

Contact information
Genevieve Konopka, Ph.D.
Department of Neuroscience, University of Texas Southwestern Medical Center,
5323 Harry Hines Blvd., ND4.300, Dallas, TX 75390-9111
TEL: 214-648-5135, FAX: 214-648-1801,
Email: Genevieve.Konopka@utsouthwestern.edu

Abstract

Cortical development is a tightly controlled process and any deviation during development may increase the susceptibility to neurodevelopmental disorders, such as autism spectrum disorders (ASD). Numerous studies identified mutations in *FOXP1*, a transcription factor enriched in the neocortex, as causal for ASD and FOXP1 syndrome. Our group has shown that *Foxp1* deletion in the mouse cortex leads to overall reduced cortex thickness, alterations in cortical lamination, and changes in the relative thickness of cortical layers. However, the developmental and cell type-specific mechanisms underlying these changes remained unclear. This work characterizes the developmental requirement of neocortical *Foxp1* at key embryonic and perinatal ages using a conditional knock-out of *Foxp1*. We find that *Foxp1* deletion results in accelerated pseudo-age during early neurogenesis, increased cell cycle exit during late neurogenesis, altered gene expression and chromatin accessibility, and selective migration deficits in a subset of upper-layer neurons. These data explain the postnatal differences observed in cortical layers and relative cortical thickness. We also highlight genes regulated by FOXP1 and their enrichment with high-confidence ASD or synaptic genes. Together, these results underscore a network of neurodevelopmental disorder-related genes that may serve as potential modulatory targets for postnatal modification relevant to ASD and FOXP1 syndrome.

Introduction

The neocortex is radially organized into layers, each of which is enriched with specialized subtypes of neurons generated from deepest to superficial layers to form its laminar structure. Excitatory glutamatergic neurons are sequentially generated from progenitors residing in two germinal compartments — the ventricular zone (VZ) and the subventricular zone (SVZ) — lining the cerebral ventricles. Apical radial glia (aRGs) in the VZ divide to produce more aRGs and to directly generate neurons (Malatesta et al. 2000; Kriegstein and Götz 2003). aRGs also indirectly generate neurons by producing intermediate progenitor cells (IPs) whose cell bodies reside in the SVZ (Götz and Huttner 2005). IPs also self-amplify and therein boost neuronal production (Kowalczyk et al. 2009; Vasistha et al. 2015; Mihalas et al. 2016). In the mouse, the generation of excitatory neurons for a specific cortical layer peaks on different embryonic days in an inside-out manner. Deep layer neurons are born first (L5-6), followed by the generation of neurons for the upper layers (L2-4) (Frantz and McConnell 1996; Shen et al. 2006; Molyneaux et al. 2007; Franco et al. 2012; Gao et al. 2014; Simon et al. 2020; Park et al. 2022).

Development of the neocortex requires the orchestrated execution of a series of crucial processes. Deviations in this fine-tuned process may increase the susceptibility to neurodevelopmental disorders (Klingler et al. 2021). FOXP1 (Forkhead Box Protein P1) is a transcription factor that plays a crucial role in various developmental processes, including cortical development. FOXP1 syndrome is a neurodevelopmental disorder caused by mutations or deletions affecting the *FOXP1* gene (Pariani et al. 2009; Hamdan et al. 2010; O'Roak et al. 2011). Individuals with FOXP1 syndrome often exhibit a range of neurological and behavioral symptoms, including intellectual disability, language impairments, motor delays, and autism spectrum disorders (ASD) (Carr et al. 2010; Meerschaut et al. 2017; Siper et al. 2017). Moreover, many individuals exhibit ADHD symptoms, such as inattention and hyperactivity (Meerschaut et al. 2017; Siper et al. 2017). However, the specific features of FOXP1 syndrome can vary widely

(Lozano et al. 2021). Previous work suggests that FOXP1 regulates other genes encoding proteins with roles in neural development (Braccioli et al. 2017; Anderson et al. 2020). However, further insight into the cellular and molecular regulators underlying these changes in the cortex is needed.

Mouse models with *Foxp1* deletion have shown phenotypes relevant for individuals with *FOXP1* variants. *Foxp1* whole-body heterozygous or brain-specific knock-out mice have reduced neonatal ultrasonic vocalization (USVs), hyperactivity, and metabolic disturbances and mitochondrial dysfunction in striatal neurons (Araujo et al. 2015; Bacon et al. 2015; Fröhlich et al. 2017; Wang et al. 2022a). Further, knocking-down *Foxp1* in mice via *in utero* electroporation during mid-neurogenesis results in broad migration deficits observed at late embryonic stages (Li et al. 2015; Braccioli et al. 2017). Additionally, removing *Foxp1* in mice from neocortical progenitors and their progeny impairs neonatal and adult USVs and results in social interaction deficits, hyperactivity, spatial learning deficits, and impairments in synaptic plasticity (Araujo et al. 2017; Usui et al. 2017). Observed neuroanatomical changes include a reduction in neocortical size and abnormal positioning of neurons in the deep layers of the postnatal mouse neocortex (Usui et al. 2017). However, the molecular changes underlying these phenotypes remained undefined.

Postnatally, FOXP1 exhibits layer-specific expression, restricted to neurons of L3-5 and sparsely in L6a (Hisaoka et al. 2010; Usui et al. 2017). During embryonic development, FOXP1 also has restricted expression in distinct cell types across neocortical development (Braccioli et al. 2017; Pearson et al. 2020). In proliferative cells outside of the nervous system, FOXP1 acts as a transcriptional repressor (Feng et al. 2010; Zhang et al. 2010). Work in embryonic stem cells shows that FOXP1 preferentially binds distinct sequence motifs compared to differentiated cells, thereby regulating the expression of pluripotent versus differentiation genes (Gabut et al. 2011).

Together, these data raise the question of whether FOXP1 may serve distinct cell type-specific roles across neocortical development.

In this study, we sought to delineate the molecular contributions of FOXP1 in a cell type-specific manner across development. We investigated the gene regulatory network governed by FOXP1 during murine neocortical development, focusing on its cell type-specific functions. *Foxp1* conditional knockouts (*Foxp1* cKOs) were generated by selectively deleting *Foxp1* from excitatory neocortical neural progenitors using *Emx1*.Cre (Gorski et al. 2002). Utilizing single-nucleus RNA sequencing (snRNA-seq) and single-nucleus Assay for Transposase-Accessible Chromatin sequencing (snATAC-seq), we ascertained the genetic and epigenetic regulatory mechanisms controlled by FOXP1. We aimed to understand the impact of FOXP1 on genetic and epigenetic mechanisms and distinctions based on cell type at various ages. Additionally, we sought to investigate whether *Foxp1* loss affects cell fate specification or migration.

We find that FOXP1 exerts opposing and congruent regulation on gene expression and chromatin accessibility during early neocortical development, particularly impacting genes associated with ASD and synaptic function. We report enrichment of FOX motifs in a cell type-specific manner, indicating genomic sites for binding and regulation by FOXP1. Thymidine analogue pulsing showed a change in cell cycle exit rates in *Foxp1* cKOs during late neurogenesis. Thymidine analogue birth-dating revealed migration deficits in an upper-layer (UL) neuron subpopulation and identified genetic players potentially underlying this phenotype. Overall, *Foxp1* loss affects genes encoding proteins enriched in ASD and the synapse, cell cycle exit, cell type-specific transcription factors, cell type proportions, and migration of specific UL neurons.

Understanding the dynamic genetic and epigenetic landscape during development is crucial for deciphering typical and atypical neural development. Our study elucidates the role of

FOXP1 at key neurogenic stages and provides insights into potential convergent disease mechanisms underlying distinct molecular etiologies.

Results

To determine the cell type-specific roles of FOXP1 across early neocortical development, we generated *Foxp1 cKO* and control littermates by crossing *Foxp1*$^{flox/flox}$ mice with mice expressing *Emx1.Cre* (Gorski et al. 2002; Feng et al. 2010). This induces recombination embryonically in excitatory neurons and progenitors derived from the dorsal telencephalon. We dissected whole cortex and used the 10X Genomics platform to perform snRNA-seq and snATAC-seq (Supplemental Fig. 1, 2).

As FOXP1 exhibits temporal regulation in distinct cell types throughout development in the neocortex, we selected three timepoints for sequencing: embryonic day (E) 13.5, E16.5, and postnatal day (P)0.5. At E13.5, FOXP1 is expressed in neuronal progenitors and post-mitotic neurons, but by E16.5 FOXP1 expression is enriched in neurons and sparsely expressed in neuronal progenitors (Supplemental Fig. 3) (Pearson et al. 2020). Postnatally FOXP1 exhibits layer-restricted expression in L3-5 and sparse expression in L6a (Hisaoka et al. 2010; Usui et al. 2017; Pearson et al. 2020).

The snRNA-seq dataset of the whole cortex ("*all cells*" dataset) captures expected cell types found in a developing neocortex such as choroid plexus, microglia, and neuronal cells (Supplemental Fig. 1). The total nuclei per timepoint consists of E13.5: 56,557; E16.5: 36,708; P0.5: 84,075 (Supplemental Fig. 1). Cell types were annotated using canonical markers and existing neocortical datasets (Loo et al. 2019; Di Bella et al. 2021) (see Methods). However, since the *Emx1-Cre* is known to recombine in the excitatory neuronal lineage, we subset only cell types from this lineage to use for downstream analyses ("excitatory lineage" dataset, Fig. 1A,B).

After we subset and re-cluster, we detected expected cell types such as aRGs, IPs, migrating neurons, and neurons from L5-6 or L2-4, which are colored by cell type in the UMAP representation (Fig. 1B). We observed 16 clusters at E13.5, 21 clusters at E16.5, and 23 at P0.5 (Supplemental Fig. 4). We collapse cell type clusters into “broad” cell type clusters (*CellType_coll*) for downstream analyses and only show these clusters for ease of visualization (Fig. 1B). Across the three timepoints, broad cell type clusters had representation from all biological replicates (Fig. 1C).

Analysis of gene expression differences showed that the *Foxp1* loss led to a notable increase in the number of upregulated differentially expressed genes (DEGs) in all cell types at embryonic timepoints and at P0.5 in only migrating neurons (Fig. 1D). FOXP1 is known to act as a transcriptional repressor, especially in cells with a proliferative capacity (Feng et al. 2010; Zhang et al. 2010). Therefore, the observed pattern of more upregulated DEGs in *Foxp1* cKOs compared to controls supports a role for FOXP1 as a transcriptional repressor during early neocortical development. However, at P0.5, L5-6 and L2-4 neurons show the opposite with more downregulated DEGs, pointing to a potential shift in the role of FOXP1 in postnatal cell types versus embryonic ones (Fig. 1D).

We also performed snATAC-seq on whole cortical tissue (Supplemental Fig. 2) and annotated the data using label transfer from our snRNA-seq dataset (see Methods). We subset the excitatory neuronal lineage to facilitate comparative analysis with our excitatory lineage snRNA-seq data (Fig. 2A). Using broad cell-type annotation (*CellType_coll*), we identified all cell types found in our snRNA-seq data, ensuring consistency across our datasets (Fig. 2B). Each cluster has representation from each biological replicate (Fig. 2C). To understand the potential regulatory role of FOXP1 in chromatin accessibility, we generated differentially accessible regions (DARs) between genotypes (see Methods, Fig. 2A,D). In contrast to our observation of more

upregulated DEGs across all embryonic timepoints and cell types, and migrating neurons at P0.5, there are more closed DARs at E13.5 and P0.5 but not at E16.5 (Fig. 2D).

These results indicate a shift in chromatin accessibility due to a *Foxp1* loss across the surveyed developmental timepoints. Changes in chromatin accessibility at the regulatory regions of a gene can influence the transcriptional activity of that gene. Overlapping DARs and DEGs indicate a possible regulatory relationship by which changes in chromatin accessibility may contribute to the corresponding changes in gene expression. For example, *Sox4* is linked to an open DAR in cKOs at E13.5 and an upregulated DEG (Supplemental Tables 1, 2). Moreover, FOXP1 may regulate other transcription factors that recruit chromatin remodeling complexes or histone-modifying enzymes, resulting in changes to chromatin structure and accessibility (Li et al. 2004; Jepsen et al. 2008; Chokas et al. 2010).

Differentially expressed genes and differentially accessible regions are enriched with ASD-relevant and synaptic genes

To gain insights into the potential disease-relevant roles of DEGs and DARs, we employed hypergeometric enrichment to assess overlap with ASD datasets (Fig. 3A,B). Specifically, we used SFARI high-confidence ASD-associated genes (Banerjee-Basu and Packer 2010), or the top ASD-associated genes identified in two large exome sequencing studies of ASD (Satterstrom et al. 2020; Fu et al. 2022). Overall, we report that DEGs and DARs across these three developmental timepoints showed enrichment with high-confidence ASD genes.

Neurons had robust association with ASD-relevant genes in terms of both DEGs and DARs. Though we anticipated enrichment between neuronal DEGs and ASD genes based on previous work in both postnatal mice and human brain organoids (Araujo et al. 2015; Park et al. 2023), we asked whether neural progenitor cells (NPCs) would also share this enrichment since this has not been previously investigated in *Foxp1* cKOs. Both aRGs and IPs also exhibit significant enrichment with ASD genes (Fig. 3A,B), albeit to a lesser extent than neurons. Other

work reported that mutations in select ASD-relevant genes altered the timing of neurogenesis of deep-layer neurons, either speeding it up or delaying it (Paulsen et al. 2022). Thus, altered gene expression in NPCs with FOXP1 loss may contribute to an increased susceptibility to neurodevelopmental disorders (NDDs).

We also noted the regulation of specific ASD relevant genes at the interface of both DEGs and DARs. Upregulated/open DAR genes in *Foxp1* cKOs include *Nr2f1*, *Cdh13*, *Celf4*, and *Srgap2* whereas downregulated/closed DAR genes include *Setbp1*, *Cdh8*, and *Sh3rf2*, all of which have been implicated in NDDs, ASD, or ID (Wilhelm et al. 2012; Katayama et al. 2016; Wang et al. 2018; Morgan et al. 2021; Cardo et al. 2023). Thus, FOXP1 may both facilitate and repress the expression of ASD relevant genes.

We next sought to investigate DEG and DAR enrichment with genes that are either pre- or post-synaptically associated (Sorokina et al. 2022). This was motivated by other work highlighting the role of ASD genes in synaptic function and neuronal activity (Rubenstein and Merzenich 2003; Gogolla et al. 2009; de la Torre-Ubieta et al. 2016; Quesnel-Vallières et al. 2019). Both DEGs and DARs display significant enrichment with presynaptic- and post-synaptic-associated genes (Fig. 3A,B; Supplemental Table 3). Interestingly, even neuronal progenitor DEGs and DARs demonstrated synaptic gene enrichment. Past work has highlighted a key role for bioelectrical properties of progenitors in neural fate determination and neuronal migration (Vitali et al. 2018). However, it is possible that these synaptically-associated genes also perform distinct functions in progenitors in addition to their known synaptic functions. Overall, these findings help underscore that genes regulated by FOXP1 are strongly associated with synaptic genes. This is important to highlight given that synaptic transmission is commonly disrupted in other ASD models (Quesnel-Vallières et al. 2019).

FOX motifs are enriched in specific cell types across development within shared DEGs and DARs

To assess the functional significance of DARs, we searched for enriched transcription factor (TF) motifs to infer potential cell type and age-specific transcriptional regulation (Fig. 3C). We identified enriched TF motifs across all timepoints in both open and closed DARs. Notably, migrating neurons exhibited the highest number of enriched motifs, while deep layer (DL) neurons had the fewest (Fig. 3C). The embryonic timepoints had more motif enrichment than the postnatal timepoint. We then specifically investigated the presence of enriched FOX family motifs (Fig. 3D). Enriched FOX motifs were observed at all timepoints, except for E13.5 in open aRG cells or L5-6 neurons, and at E16.5 and P0.5 in closed L5-6 neurons (Fig. 3D).

We next asked if motif enrichment was linked to gene expression. We generated a list of overlapping DEGs with associated DARs, identifying numerous overlaps across development in specific cell types. Then, we assessed enriched FOX motifs within these overlaps by specific cell type (Fig. 3E). To determine genes potentially directly bound and regulated by FOXP1, we noted both enriched FOXP1 and general FOX motifs among these comparisons (Fig. 3E, FOX denoted by a solid star symbol, FOXP1 motifs by a striped star symbol). Either FOXP1 or FOX motifs are enriched in most comparisons. One exception is L5-6 neurons at all ages, cell types where we expect little to no FOXP1 expression. Another exception is L2-4 neurons at P0.5 (Fig. 3E). However, we noted enrichment of other transcription factors by cell type that may work together with or downstream of FOXP1 to regulate gene expression. We detected upregulated DEGs and open DARs in aRGs for *Yy1* (*Yin Yang 1*) and *Nr2f1/COUP-TFI* together with enriched motifs for these genes among additional DEGs/DARs. A previous study suggested that *Foxp* genes interact with *Nr2f1* and *Yy1 in vitro* and both have implications for neurodevelopmental disorders and brain development (Estruch et al. 2018). These findings suggest FOXP1 may dynamically regulate gene expression in specific cell types via both directly binding to DNA and/or indirect ways.

Finally, we wanted to determine patterns across development for cell type-specific regulation by FOXP1. We found higher overlap in closed DARs and downregulated DEGs at E13.5 and P0.5 with the inverse at E16.5, highlighting dynamic regulation by FOXP1. Overall, many of the DEGs and DARs include transcription factors, suggesting that FOXP1 regulates a large network of genes that themselves exert various regulatory effects (Supplemental Tables 1,2).

Loss of Foxp1 results in a more mature pseudo-developmental trajectory at E13.5

We next specifically focused on the role of FOXP1 at E13.5, when FOXP1 is expressed in both neuronal progenitors and post-mitotic neurons (Supplemental Fig. 3). We sought to determine how FOXP1 differentially regulates the developmental trajectory of each of these cellular populations. Using only the E13.5 snRNA-seq data, we use broad cluster per cell type for ease of visualization (*CellType_coll*) (Fig. 4A). We then performed pseudotime analysis using diffusion pseudotime (see Methods) (Angerer et al. 2016). We report that pseudotime distribution of cell types in *Foxp1* cKOs is condensed compared to the control path, with reduced separation between cell transitions, especially between aRGs and IPs (Fig. 4B, Supplemental Fig. 5A). Overall, the pseudo-age value for all cKO cell types is older than that assigned to control cells (Fig. 4B, Supplemental Fig. 5A). The total path length of the two developmental trajectories is significantly different (Mann-Whitney test with continuity correction adj. p-val < 2.2e-16). Therefore, FOXP1 loss leads to an estimated older cellular pseudo-age, indicating that FOXP1 plays a role regulating the development of cells during early neurogenesis.

Loss of Foxp1 alters gene expression to reduce progenitor maintenance and promote differentiation

To determine the signaling pathways downstream of FOXP1 driving changes in the developmental trajectories of cell types, we assessed the functional enrichment of gene ontology (GO) categories within the DEGs/DARs of aRGs (Supplemental Fig. 5C,D; Supplemental Tables 5, 6). Upregulated DEGs had GO terms such as axon guidance, semaphorin receptor activity,

negative regulation of glial cell proliferation, and positive regulation of cell morphogenesis involved in differentiation (Supplemental Fig. 5C, Supplemental Table 5). GO terms for open DARs in aRGs also represented positive regulation of cell differentiation and glial cell differentiation, chromatin remodeling, Wnt signaling, positive regulation of gene expression, and regulation of microtubule (de)polymerization (Supplemental Fig. 5D, Supplemental Table 6). Conversely, downregulated DEGs had GO terms for SUMO transferase activity, ubiquitin protein ligase activity, and microtubule binding (Supplemental Fig. 5C, *right*). Closed DAR Go terms involved regulation of cell migration, transcription co-repressor activity, ephrin receptor activity, and transmembrane receptor protein tyrosine kinase activity (Supplemental Fig. 5D, *bottom*; Supplemental Table 6). These terms highlight changes in the biological function of aRGs and the underlying mechanisms driving observed changes in developmental trajectories. These results suggest that FOXP1 represses genes associated with differentiation in aRGs to permit proper maintenance of self-proliferation. Thus, at E13.5, aRGs in cKOs express more pro-differentiation genes, resulting in an advanced estimated pseudo-age.

snRNA-seq data can not only reveal changes in differential expression but also changes in cell-type proportions. At E13.5, we observed a notable shift in cell-type proportions, with significantly fewer aRGs (adj. p-val <2.97e-06) together with a significant relative increase in L5-6 neurons in cKOs (adj. p-value < 2.2e-16, Fig. 4C). Other work reported an increased prevalence of neurogenic-type divisions in radial glia of mice lacking *Foxp1* (Pearson et al. 2020). Together, the changes in gene expression and cell type proportions at E13.5 indicate a promotion of neuronal differentiation with FOXP1 loss.

S-Phase length is increased in progenitors upon loss of Foxp1

To assess the mechanism of improper maintenance of the progenitor pool, we use two distinguishable thymidine analogues, BrdU (5-bromo-2'-deoxyuridine) and EdU (5-Ethynyl-2′-deoxyuridine), to estimate the length of S-Phase at E13.5 (Martynoga et al. 2005). By co-staining

with KI-67, a marker of proliferating cells, we find that neural progenitors in *Foxp1* cKOs have a significantly lengthened S-Phase (Fig. 4D) and cell cycle length (Supplemental Fig. 5B). Prior work has shown increased S-Phase length of apical and basal progenitors in the mouse cortex with sustained *Pax6* expression (Wong et al. 2015). *Foxg1* is a regulator of neurogenic commitment (Hanashima et al. 2002; Hettige et al. 2022; Wang et al. 2022b) and is upregulated in neural progenitors at E13.5 in *Foxp1* cKOs (Supplemental Tables 1, 2). *Yy1*, a gene involved in progenitor maintenance (Knauss et al. 2018), is also upregulated in DEGs / open DARs in aRGs of cKOs (Supplemental Tables 1, 2). Finally, *Cdk14*, a downregulated DEG and closed DAR, promotes cell proliferation (Chen et al. 2019).

To determine whether other aspects of the cell cycle are impacted by FOXP1 loss, we first used phosphorylated histone M3 (PH3) staining as a proxy of M-Phase. We detected no significant differences in PH3+ cells between cKOs and controls at E13.5 (Supplemental Fig. 5F). Downregulated DEGs in aRGs are enriched for GO terms related to protein ubiquitination (Supplemental Fig. 5C, right). Although apoptotic pathways are known to be tightly regulated by post-translational modifications such as ubiquitination (Roberts et al. 2022), we did not detect a change in apoptosis as measured by cleaved-caspase-3 between genotypes at E13.5 (Supplemental Fig. 5G). Thus, the progenitor pool is altered through fewer aRGs in cKO vs controls (Fig. 4C), a longer S-Phase, and dysregulation in genes necessary for progenitor maintenance.

Precocious generation of DL neurons

The change in pseudo-development trajectory and enriched GO terms for differentiation in aRGs prompted further investigation into the shifting neurogenic potential of cKO aRGs. Our previous work demonstrated cyto-architectural changes in the P7 cortex of *Foxp1* cKOs. We observed a relatively thinner cortex overall with a reduction in thickness of the DL but a relatively thicker L2-4 (Usui et al. 2017). At E13.5, we find that *Foxp1* cKOs have a relatively thicker TBR1+ cell layer

(Fig. 4E). However, this significant difference in TBR1+ relative thickness is not detected at E14.5 or E15.5 and reverses at E16.5 when *Foxp1* cKOs have a relatively thinner TBR1+ layer compared to controls (Fig. 4E). These results may be attributed to an increased frequency of differentiative divisions of neural progenitors at E13.5 in *Foxp1* cKOs (Pearson et al. 2020). Increased neurogenic division may lead to a depletion of the progenitor pool and then to an overall reduction in neuron production observed postnatally in cKOs. In support of this mechanism, at E13.5, cKO L5-6 neurons exhibit upregulated expression of DL markers *Bcl11b*, *Tshz2*, and general neuronal marker *Rbfox3 / Neun* (Fig. 4F). Thus, *Foxp1* cKO mice exhibit changes in DL gene expression patterns that may underlie observed postnatal phenotypes.

Accelerated cell cycle exit from E15.5 to E16.5 with loss of FOXP1

Corticogenesis is characterized by the slowing down of the cell cycle in neural progenitor cells and an increasing frequency of differentiative divisions, resulting in an increase in the rate of neurogenesis (Borrell and Calegari 2014). In *Foxp1* cKOs, corticogenesis appears initially sped up (Fig. 4B). We asked whether the reversal of TBR1+ layer thickness, a marker for DL neurons, from E13.5 to E16.5 could be due to changes in neurogenesis, or the rate of cell cycle exit at one of the intermediate time-points. To measure changes in cell cycle exit, we pulsed EdU once at E15.5 and analyzed changes after 24 hours. EdU is incorporated into cells undergoing DNA-replication during S-phase. We quantified cells that were cycling at the time of injection (EdU+) but are no longer proliferative 24 hours later (KI67-), indicating they have left the cell cycle (EdU+Ki67- / EdU cells) (Fig. 5A) (Yoon et al. 2017). We observe that cKOs have more cells that have exited the cell cycle than controls (Fig. 5A). Interestingly, the faster cell-cycle exit of NPCs results in a bias toward neuron production at the expense of self-renewing divisions in NPCs. This may point to a reason behind the relatively thicker L2-4 reported in *Foxp1* cKOs at P7 (despite an overall thinner neocortex) since most of the neurons born at E16.5 acquire UL identities

(Molyneaux et al. 2007; Usui et al. 2017). In fact, we observe a relative increase in L2-4 neuron proportions in cKOs by P0.5 in our snRNA-seq data (Fig. 5B *left*), but not at E16.5 (Fig. 5B *right*).

Dysregulation of cell type-specific markers in E16.5 neurons

Canonically, UL neurons are born from mid-to-late-corticogenesis, with newborn neurons at E16.5 overwhelmingly attaining UL identities (Molyneaux et al. 2007; Telley et al. 2019). Previous work has identified a growing body of layer- and neuronal subtype-specific genes. Some select layer-specific genes include *Tbr1*/*Foxp2*/*Sox5* for layer 6 neurons, *Fezf2* or *Bcl11b/Ctip2* for layer 5 neurons, *Cux1*/*Cux2* for UL neurons with *Rorb* defining layer 4 and *Pou3f2* layer 2/3 neurons (Molyneaux et al. 2007). At E16.5 in *Foxp1* cKOs, UL neurons have downregulated expression of canonical L2-4 transcription factors *Cux1*, *Cux2*, *Rorb,* and *Bcl6* yet upregulated expression of *Neurod6*, a general excitatory marker (Fig. 5C). In contrast, *Nfi-* (*Nuclear factor I*) family members *Nfia* / *Nfib* / *Nfix* exhibit increased expression in UL neurons at E16.5 (Fig. 5D). These genes are relevant for proliferation and differentiation of neural progenitor cells and proper migration of neurons (Heng et al. 2012; Martynoga et al. 2013; Betancourt et al. 2014). Though developmental and migration defects are most often seen in knock-down of *Nfi-* factors, misexpression or changes in timing of expression may result in changes in cell type identity (Bunt et al. 2017). DL marker expression is also altered as evidenced by downregulation of *Bcl11b*/*Ctip2* and upregulation of the excitatory neuron marker *Neurod6* in DL neurons at E16.5 in cKOs (Fig. 5E. Together, these data indicate a role for FOXP1 in facilitating the expression of canonical L2-4 genes by refining the identity of both UL and DL neurons at E16.5 gene expression and thus cell identity.

Neuronal migration deficits with loss of FOXP1

Previous work has reported that knock-down of *Foxp1* via *in utero* electroporation results in broad neuronal migration deficits (Li et al. 2015; Braccioli et al. 2017). However, genetic knock-down of *Foxp1* does not readily reflect such a stark phenotype, but rather a more restricted change, in the

postnatal neocortex with mispositioning of CUX1+ cells, a traditional marker of L2-4 neurons, in the DLs of *Foxp1* cKOs at P7 (Usui et al. 2017). We asked whether these ectopic CUX1+ cells were a result of migration deficits or of changes in cell fate specification. We reason that since FOXP1 is also present in neuronal progenitors, loss of *Foxp1* may affect neuronal fate. To answer this, we pulsed the thymidine analogue EdU at discrete embryonic time-points and analyzed brains at P7. We co-stained brain sections with CUX1 (L2-4 marker), CTIP2 (L5-6 marker), and EdU to identify EdU+CUX1+ cells in the DL as marked by the start of CTIP2+ cell layer (Fig. 6A). We find that EdU injections at E15.5 result in EdU+CUX1+ cells in the DLs (Fig. 6A,B). We do not observe this phenotype at earlier time-points (data not shown). The majority of neurons generated at E15.5 are fated to form UL neurons (Molyneaux et al. 2007). Our results suggest that UL neurons are being generated at the anticipated times but are unable to migrate to their expected destination in the *Foxp1* cKOs. We quantified the number of EdU+ cells in the UL and DL and determined that *Foxp1* cKOs have a significantly higher number of EdU+ cells in the DLs compared to controls (Fig. 6C). The number of EdU+ cells in the DLs further suggests a migration deficit phenotype.

We next assessed these effects at P16.5 to determine whether migration was significantly delayed in *Foxp1* cKOs. We still observed EdU+/CTIP2- cells in the DL (Fig. 6D,E, Supplemental Fig. 6A,B). As the number of EdU+ cells in the DLs remains significantly higher in cKOs this indicates a primary migration deficit, not a delay in neuronal migration (Fig. 6E). Further, DEGs and DARs at E16.5 and P0.5 also highlight genes with key functions in neuronal migration, further supporting dysregulation in migration cues (Supplemental Fig. 6C, Supplemental Table 5, 6).

Discussion

Here, we highlighted cell type-specific gene expression and chromatin accessibility changes that occur at developmental timepoints when FOXP1 expression is lost. At E13.5, FOXP1 is expressed in both neural progenitors and post-mitotic neurons. Upon *Foxp1* loss in neural

progenitors we observe advanced pseudo-age leading to faster generation of DL neurons (Fig. 7A). These modified neural progenitors have altered expression of genes (e.g., *Cdk14, Sox4 / Sox11 or Yy1*) that may underlie their increased neurogenic divisions. However, the precocious development of DL neurons reduced the progenitor population, leading to reduced relative proportions of DL neurons (Fig. 7A). We also find that by E16.5 neural progenitors also exit the cell cycle faster. Neurons are also affected by the loss of FOXP1 with impacts on migration and expression of canonical layer-specific TFs. Finally, all of these cellular phenotypes occur in parallel with molecular changes in gene expression relevant to synaptic function and NDD pathophysiology (Fig. 7B).

Regulation of molecular identity of cell types

Foxp1 has an important role in refining molecular identity of cell types, as shown by the detection of canonical layer-specific TFs as DEGs or DARs. For example, cKO DL neurons at E13.5 had elevated expression of DL TFs (*Tshz2*, *Bcl11b,* Supplemental Table 1). At E16.5, *Bcl11b* is a downregulated DEG in DL cKO neurons (Supplemental Table 1). At P0.5, cKO DL neurons had downregulated expression of *Tshz2* (Supplemental Table 1). In UL neurons, at E16.5, the expression of L2-4 marker genes is downregulated in cKOs (e.g., *Cux1* and *Bcl6*) and this is reversed by P0.5 (Supplemental Table 1). These changes highlight a requirement for FOXP1 in fine-tuning the appropriate molecular signatures of cell types across development.

Selective migration deficits of a subtype of UL neuron

Previous work has highlighted the importance of *Foxp1* in regulating neuronal migration (Li et al. 2015; Braccioli et al. 2017; Usui et al. 2017). Our method of birth-dating neurons in *Foxp1* cKO animals with a thymidine-analogue allowed us to definitively disentangle a migration impairment from a cell fate specification phenotype. We opted to use a thymidine analogue pulse, although other techniques exist, such as FlashTag (Govindan et al. 2018), which tag dividing progenitors that contact the ventricular surface to mark a more precise group of cells. Thymidine analogues

incorporate their mark into DNA of all progenitors and their progeny during DNA replication in S-Phase. Though the cells marked using this method are broader, we could observe patterns of cell generation based on date of injection and track neuronal generation and the ultimate position of neurons based on birthdate.

Using single cell genomics, we also gained unbiased genomic profiling of cells at each timepoint (Fig. 7B). At E16.5, we observed that *Nfia* and *Nfib* are upregulated in L2-4 neurons (Fig. 5D, Supplemental Table 1), but by P0.5, *Nfia* is a downregulated DEG and closed DAR in migrating neurons (Supplemental Table 1,2). *Nfi* family factors are also downregulated DEGs at P0.5 in UL neurons (Supplemental Table 1). These factors have roles in regulating neuronal migration and neural development (Martynoga et al. 2013; Paul et al. 2014; Bunt et al. 2017; Li et al. 2017), emphasizing key potential players in a FOXP1 regulatory network that affect migration upon disruption. Previous work showed that FOXP2 acts with NFI cofactors in a DNA-binding-independent manner to drive maturation of excitatory cortical neurons (Hickey et al. 2019), suggesting a potentially similar role for FOXP1 in UL excitatory neurons where Foxp2 is not expressed.

Other genes of interest relevant for regulation of migration include *Scrt2* (*Scratch2*), which is a downregulated DEG and closed DAR (Supplemental Table 1,2). SCRT2 has modulatory roles in cortical neurogenesis and migration, as well as a role in the delay of the radial migration of neurons, especially UL neurons (Paul et al. 2014). Semaphorins and their receptors, plexins which are guidance molecules and receptors involved in regulating critical patterning (Limoni and Niquille 2021), are also affected with loss of FOXP1. At E16.5, *Plxna4* (*Plexin gene PlexinA4*) and *Nav2* are downregulated DEGs in cKO L2-4 neurons and migrating neurons (Supplemental Table 1). Additionally, *Plxna4* is downregulated in *Foxp1* cKO neurons at P0.5 in DL neurons (Supplemental Table 1). PLXNA4 is a co-receptor with NRP1 for SEMA3A, and knocking-down *PlxnA4* resulted in radial migration defects (Chen et al. 2008). In our snATAC-seq data, we found

Foxo6 associated with a closed DAR in cKOs at P0.5 (Supplemental Table 2). FOXO6 directly binds to PLXNA4 to facilitate SEMA3A signaling which is important for radial migration during late corticogenesis (Chen et al. 2008; Limoni et al. 2021; Limoni and Niquille 2021). *Foxo6*$^{+/-}$ mice show radial migration deficits, with neurons halted in the deep layers of the developing cortex, and a reduction in PLXNA4 levels (Paap et al. 2016). Thus, reduced expression of *Foxo6* and *PlxnA4* may underlie the selective neuronal migration deficit we observe in *Foxp1* cKOs.

Foxp1 regulates genes relevant for ASD

Foxp1 is a high confidence ASD-associated gene and previous work has shown that FOXP1 regulates genes important for neurodevelopment (Araujo et al. 2015; Usui et al. 2017; Satterstrom et al. 2020; Fu et al. 2022). We specifically highlight increased enrichment of DEGs and DARs with ASD genes across all timepoints. Given the heterogeneity of genetic mutations and molecular mechanisms contributing to ASD, it is important to understand convergent mechanisms for ASD associated genes. For example, past work seeking to coalesce changes due to ASD risk genes found asynchronous maturation of DL neurons and interneurons (Paulsen et al. 2022). Here, we report changes in molecular function and generation of DL neurons and additional changes in UL neurons (Fig. 7B). Thus, these data may be helpful as the field continues to gain additional insight to determine shared changes in UL neurons in other models assessing ASD risk genes.

Foxp1 regulates genes encoding synaptic proteins across development

We also reported that both DEGs and DARs are enriched for synaptic protein encoding genes, especially in postnatal tissue. Though synaptic protein relevant gene expression changes are present from early development, their persistence and increasing enrichment among DEGs over development suggest that these changes are most relevant postnatally. This may implicate changes in micro- and macro- circuity of the developing neocortex. Recent work has highlighted the importance of *Foxp1* in regulating potassium currents in striatal neurons (Khandelwal et al.

2021). This underscores the possibility of *Foxp1* serving a similar role in excitatory cortical neurons.

Conclusions

Our data provide cellular resolution insights into the *in vivo* roles of *Foxp1* across early neocortical development. We show that *Foxp1* is important in a cell type-specific manner for the function of neural progenitor cells, impacting neurogenesis rates indirectly by altering their proliferative function. We also detail how FOXP1 deletion directly impacts cortical migration. Moreover, we identify the differentially expressed genes and differentially accessible regions due to *Foxp1* loss that are enriched for genes encoding ASD and synaptic proteins and may underlie the cell type-specific phenotypes. These findings provide opportunities for cellular and molecular manipulations to rescue FOXP1 syndrome and ASD-relevant phenotypes.

Methods

Mice

All procedures were approved by the Institutional Animal Care and Use Committee of UT Southwestern. *Emx1-Cre (Gorski et al. 2002)* (#005628, Jackson Laboratory); *Foxp1*$^{flox/flox}$ mice (Feng et al. 2010), provided by Dr. Haley Tucker, were backcrossed to C57BL/6J for at least 10 generations to obtain congenic animals as previously described (Araujo et al. 2017; Usui et al. 2017). We bred *Emx1-Cre/Foxp1*$^{flox/flox}$ males to *Foxp1*$^{flox/flox}$ female mice to obtain *Foxp1* cKO and control mice. Mice were group-housed under a 12 h light/dark cycle and given ad libitum access to food and water. Mice of both sexes were used for all experiments. For timed breeding, female mice were paired one night with male mouse and noon of the following day is considered E0.5 (Usui et al. 2017).

Single nucleus cortical tissue processing for snRNA-seq

For embryonic timepoints, tail snips from embryos were taken for genotyping while in ice-cold PBS; then whole cortex hemispheres were rapidly dissected into separate micro-centrifuge tubes, flash frozen in liquid nitrogen, and stored at -80°C. Nuclei were isolated per modifications of other isolation protocols (Habib et al. 2017; Ayhan et al. 2021a). Single nuclei were obtained from cortex samples in 1 mL ice cold Nuclei EZ lysis buffer (Sigma, #EZ PREP NUC-101), homogenized using 2 mL glass Dounce tissue grinders on ice; 15 times with pestle A and then 15 times with pestle B. The solution was then transferred to a new 2 mL microcentrifuge tube, 1 mL nuclei EZ lysis buffer was added, followed by 5-minute incubation on ice, and then centrifuged at 500xg for 5 minutes at 4°C. Supernatants were removed, pellets were resuspended in 1 mL nuclei EZ lysis buffer, followed by an additional 5-minute incubation on ice, and centrifuged at 500xg for 5 minutes at 4°C. The supernatant was then discarded, the pellet resuspended in NSB (1X PBS w/ 0.2 U/uL RNase inhibitors / 1% Ultra-pure BSA) and filtered through a 40 um Flowmi Cell Strainer (Bel-Art, H13680-0040). Nuclei were then counted on a hemocytometer (iChip) in a 10 uL of

suspension with 1:1 0.4% Trypan Blue. Nuclei concentration was adjusted for targeted sequencing of 7,000-10,000 nuclei/sample using the 10X Genomics Single Cell 3′ Reagent Kits v3 protocol to prepare libraries. 4 mice/genotype and timepoint were used, including males and females. Libraries were sequenced using the Illumina NovaSeq 6000 via the McDermott Sequencing Core at UT Southwestern.

Single nucleus cortical tissue processing for snATAC-seq.

Nuclei were isolated per modification of prior protocols (Habib et al. 2017; Ayhan et al. 2021a). Cortex samples were placed in a 2 mL Dounce homogenizer on ice with 1 mL ice cold Nuclei EZ lysis buffer (Sigma, #EZ PREP NUC-101) and homogenized 15-20 times with pestle A and then 15-20 times with pestle B. The solution was transferred to a clean 2 mL microcentrifuge tube; 1 mL nuclei EZ lysis buffer was added, followed by 4-minute incubation on ice, and centrifugation at 500xg (rcf) for 5 minutes at 4°C. The supernatant was then discarded, pellet resuspended in 300 uL-1 mL Nuclei Wash Buffer (Tris-HCl pH7.4 10 mM, NaCl 10 mM, MgCl2 3 mM, BSA 10%, Tween-20 1%), then centrifuged 500xg for 5 min at 4°C. The supernatant was discarded, and pellet resuspended in 100uL chilled diluted Nuclei Buffer on ice to allow for natural dissociation for 3-4 min. Then the final volume was brought to 300 uL-1 mL. This was then filtered through a 40um pre-wetted Flowmi Cell Strainer into a 1.5 mL microcentrifuge tube. Nuclei were then counted on a hemocytometer (iChip) in a 10 uL of suspension with 1:1 0.4% Trypan Blue. Nuclei concentration was adjusted for targeted sequencing of 4,000-7,000 nuclei/sample using the 10X Genomics Chromium Single Cell ATAC v1.1-v2 protocol to prepare libraries. 2-3 mice/genotype and timepoint were used, using both males and females. Libraries were sequenced using the McDermott Sequencing Core at UT Southwestern.

BrdU or EdU experiments

To label dividing cells and measure migration, timed pregnant *Foxp1*$^{flox/flox}$ dams were intraperitoneally injected at specified embryonic days (E12-E17) with a 50 mg/kg dose of 5-

Bromo-2'-deoxyuridine (BrdU, Sigma-Aldrich B5002) per gram of body weight of a 20 mg/mL BrdU solution dissolved in PBS or with EdU (Invitrogen, C10340) at 5 micrograms per gram body weight. For migration assays, brains were collected from injection date until P7 or P16.5. For S-Phase experiments, EdU was injected into pregnant dams, then 1.5 hours later BrdU was injected (Ti), then embryos were collected 2 hours after the first injection (30 min post-BrdU). S-Phase length (Ts) = Ti x (EdU+BrdU+ cells/ EdU+BrdU- cells); and cell cycle length (Tc) was calculated by Ts / (EdU+BrdU+ cells/ Ki67+ cells) as Ki67 marks all cycling progenitors (Martynoga et al. 2005).

All BrdU treated brains underwent antigen retrieval. For migration assays, sections were stained with rat anti-BrdU (BU1/75 (ICR1)) or mouse anti-BrdU (G3G4 (anti-BrdU), DSHB AB_2314035) and layer-specific antibody markers. For EdU treated brain, Click-iT EdU kit instructions were followed (Invitrogen, C10340). Images were acquired at 20X or 63X using a Zeiss LSM 880 confocal microscope. Stitched maximum intensity projection images are used for cell counting using FIJI.

Immunohistochemistry

For embryonic timepoints, dams were anesthetized with CO_2 and embryos were extracted rapidly and placed in a Petri dish with ice-cold 1X PBS. Tail snips were taken for genotyping. Embryos were transferred to clean Petri dishes with ice-cold 1X PBS twice before drop fixing in 4% PFA at 4°C for 30 min-overnight. Embryos were cryoprotected in 20% and 30% sucrose with 0.01% sodium azide until they sunk before mounting in OCT media and freezing on dry ice. Sections were made at 12-30 um on a Leica CM1950 Cryostat, directly mounted onto slides, and allowed to dry before processing. Neonatal and adult mice were anesthetized then transcardially perfused with 1x PBS and 4% PFA and their brains postfixed overnight in 4% PFA at 4°C. After cryoprotection in 20% and 30% sucrose with 0.01% sodium azide overnight at 4°C, brains were

embedded in Tissue-Tek CRYO-OCT Compound (#14-373-65, Thermo Fisher Scientific) and cryosectioned at 15-30 um.

For staining, all washes were done with TBS or 0.4% Triton X-100 in TBS (TBS-T) unless otherwise stated. When specified, antigen retrieval was performed in citrate buffer (10 mM tri-sodium citrate, 0.05% Tween-20, pH 6) for 10 min at 95°C. Free aldehydes were quenched with 0.3 M glycine in TBS-T for 1 hour at room temperature (RT), and sections blocked in 10% normal donkey serum (NDS) and 3% bovine serum albumin (BSA) in TBS. Sections were incubated 1.5 hours at RT or overnight at 4°C in primary antibodies diluted in 10% NDS and 3% BSA in TBS. Secondary antibody incubations were performed for 1 hour at RT or overnight at 4°C in 10% BSA in TBS. Sections were then washed three times with TBS. Sections were incubated with Hoechst or DAPI 1:2k for 2 min, then washed in TBS. Coverslips were mounted using ProLong Diamond Antifade Mountant without DAPI (#P36970, Thermo Fisher).

Antibodies include: rat anti-BrdU (BU1/75 (ICR1, 1:500)) or ms anti-BrdU (Developmental Studies Hybridoma Bank, 1:100) with antigen retrieval; rat-anti-Ki-67 (SolA15) eBioscience (14-5698-82, 1:500); rabbit anti-Foxp1 (CST 2005S, 1:500), rb anti-SOX2 (Millipore AB5603, 1:500), ms-anti-PAX6 (sc-81649, 1:200), rat-anti-EOMES (Dan11mag) (Thermo Fisher, 14-4875; 1:800), rat-anti-CTIP2 (abcam ab18465, 1:100), mouse anti-neuronal classIII ß-tubulin (Tuj1, MMS-435P, Covance, 1:800), rb-anti-CUX1 (Proteintech 117331, 1:600); rb-anti-Phospho-Histone H3 (Ser10) (PHH3; CST 9701, 1:200), and rb-anti-activated Cleaved Caspase-3 (Asp175) (CST #9661, 1:200). Species-specific secondary antibodies produced in donkey and conjugated to Alexa Fluor 488, Alexa Fluor 555, or Alexa Fluor 647 were used (Thermo Fisher Scientific, 1:1000).

Imaging and image analysis

Images were acquired using a Zeiss LSM 880 confocal microscope at the UT Southwestern Neuroscience Microscopy Facility and processed and analyzed using Zeiss ZEN Lite and FIJI. For quantifications, tile scan Z-stack images of the region of interest were acquired at 20X

magnification from similar coronal sections across at least 2–3 mice/genotype across at least 2 litters. Stitched maximum intensity projection images are used. Cell counting was done using the FIJI Cell Counter plugin. Layer thickness was measured in FIJI and normalized to total cortical thickness based on DNA stain DAPI or Hoechst. Experimenters were blind to genotype during quantification.

Migration assays quantification

Pregnant dams were injected with EdU (5 ug/g) at embryonic date of interest to label cells in S-Phase for migration. Offspring were then collected postnatally at P7.5 or P16.5. Antibodies for EdU, CUX1, and CTIP2 were used to denote cells proliferating at the time of injection, and markers of UL and DL cells respectively. Location in cortical layers for cells co-positive for EdU+CUX1+ were binned for UL and DL based anatomically on marker genes CUX1, CTIP2, and Hoechst staining. Specifically, to denote migration defects, EdU+CUX1+ cells in L5-6 were measured. Then overall locations of EdU+ cells in cortical layers were semi-automatically quantified using FIJI and grouped into UL or DL bins and compared using the Mann-Whitney test.

Cell cycle exit assay quantification

Pregnant dams were IP injected with EdU and brains from embryos from the pregnant dams were collected 24 hours later. Co-staining of coronal sections 12-25um thick with Ki-67 and EdU was used to assess the state of cells 24 hours post injection, e.g., whether they remained proliferative, KI67+ (Ki67 is a proliferating cell marker that is expressed in all mitotic cells). Cells that were EdU+Ki67- at the time of collection were cells that exited the cell cycle, while EdU+Ki67+ cells were still in the cell cycle; these were both normalized to total EdU cells. Analyses were done in FIJI and Mann-Whitney test done in GraphPad Prism.

Bioinformatics

Pre-processing of snRNA-seq data

Unprocessed sequencing data were acquired from the McDermott Sequencing Core at UT Southwestern as binary base call (BCL) files. BCL files were de-multiplexed with the 10X Genomics i7 index (used during library preparation) using Illumina's bcl2fastq v2.17.1.14 (Andrews 2010) and *mkfastq* command from 10X Genomics Cell Ranger v2.1.1 tools (Zheng et al. 2017). Extracted paired-end fastq files were checked for read quality using FASTQC v0.11.5 (Andrews 2010). The raw gene-count matrix containing cells as rows and genes as columns was used for ambient RNA removal. Ambient RNA contamination was assessed using CellBender 0.2.0 (remove-background, default parameters), a package in Python 3.7 (Fleming et al. 2023). The CellBender processed count matrix was used for downstream analyses.

Doublet removal

Scrublet and DoubletFinder (post-clustering) were used to predict doublets in the snRNA-seq data. Doublets identified within both packages were deemed as "true" doublets and removed from downstream analyses.

Intronic read ratio

As an additional measure to remove ambient RNA contamination, we used the intronic read ratio because a low intronic read ratio indicates the presence of non-nuclear transcripts (Caglayan et al. 2022). We measure intronic read ratio per cluster per sample, and clusters with an intronic read ratio below 0.5 were removed from downstream processing.

Clustering analysis

The described single-nuclei RNA-seq count matrix was used for clustering using the Seurat analysis pipeline in R (Butler et al. 2018). A Seurat object was created using default parameters. Then, nuclei with more than 10,000 molecules (*nUMI* per cell) and nuclei with more than 5% mitochondrial or ribosomal content, or fewer than 300 detected genes (*nFeature*) per nuclei at E13.5 or <400 at E16.5 and P0.5 were filtered out to discard potential doublets and low-quality

nuclei. Genes from mitochondrial chromosome and chromosomes X and Y were also removed as samples were from mixed sexes.

The filtered data were then log-normalized and scaled using a factor of 10,000 using *NormalizeData*, and regressed for covariates such as number of UMI per cells, percent mitochondrial content per cell, and batch, as described in the Seurat analysis pipeline (Butler et al. 2018). To further identify the top variable genes, the data were used to calculate principal components (PCs). Using the elbow plot and jackstraw analyses, statistically significant PCs were used to identify clusters within the data using the original Louvain algorithm. Harmony was used to regress for batch effects using *RunHarmony*. The clusters were then visualized with uniform manifold approximation and projection (UMAP) in two dimensions using reduction = "*harmony*" (Butler et al. 2018).

Cell type annotation

Cluster marker genes were identified using *FindAllMarkers* with default parameters. Our cluster markers were compared with markers from published scRNA-seq datasets from embryonic and neonatal mouse cortex (Loo et al. 2019; Di Bella et al. 2021). Cell types were assigned to clusters based on statistically significant enrichment of gene sets using the hypergeometric test (with a background of the total number of expressed genes. Additionally, clusters that overlapped significantly with multiple cell types were called for the most significant overlap (smallest adj. p value) and analyzed for expression of top marker genes of known cell types. This was combined with the cell type identities assigned when using the Seurat *LabelTransfer* protocol with the reference dataset. This is the "all-cells" (*CellType1*, *CellType1_coll*) dataset. Thereafter, only cells of the excitatory neuronal lineage were selected to continue the analysis. Total cell counts were as follows E13.5: 56,557; E16.5: 36,708; P0.5: 84,075.

Cell type re-clustering

Cells from clusters that fell into excitatory neuronal lineage categories were taken as a subset from the “all-cells” dataset and re-clustered to define cell subtypes. Clusters that were then poorly clustered or had inhibitory gene markers were excluded from further analysis. Processing was done again as described above and cell types assigned based on canonical markers and based on statistically significant enrichment from existing datasets (Loo et al. 2019; Di Bella et al. 2021) gene sets using the hypergeometric test (with a background of the total number of expressed genes: 30,578). This dataset is referred to as excitatory neuronal lineage (*CellType*, *CellType_coll*). Total cells per timepoint are E13.5: 31,543; E16.5: 26,444; P0.5: 50,672.

Differentially expressed gene (DEG) analyses

To determine differentially expressed genes, pairwise differential gene expression analysis tests were performed using MAST-GLM (Finak et al. 2015) within each broad cell type-pair between control and cKO as previously done (Ayhan et al. 2021b). Thus, only one broad cluster (e.g., apical radial glia) per major cell type was used for DEG analysis (labeled *CellType_coll*). Only DEGs with an adjusted p-value < 0.05, $|\log FC| > 0.15$ when compared to control were deemed significant (Finak et al. 2015).

DEG gene ontology (GO) analysis

Gene ontology (GO) analysis was performed in R using enrichGO via clusterProfiler with biological processes or molecular function (Wu et al. 2021). Terms were deemed significant if they had a corrected p-value < 0.05 and contained a minimum of three genes.

Overlap with outside gene databases

High-confidence ASD associated genes were downloaded from the SFARI-ASD (Banerjee-Basu and Packer 2010) database. ASD risk genes identified by two prior studies were used (Satterstrom et al. 2020; Fu et al. 2022). Synaptic genes were downloaded from Synaptome.db (Sorokina et al. 2022). Significant overlap was determined using the hypergeometric overlap test.

Pseudotime analysis

Pseudotime analysis was performed using destiny (Angerer et al. 2016) in R. Data were processed as described above. For ease of visualization, within cell-type clusters (e.g., all L2-4 cells) were collapsed (*CellType_coll*), then separated by genotype per timepoint for within-timepoint, or between-genotype comparison. For each timepoint, the control and cKO object were used for creation of the diffusion map and pseudotime required for destiny. The values of pseudotime per cell and the significant difference between pseudotime estimates was tested using the Mann Whitney test using continuity correction in R.

snATAC-seq data

Preprocessing

BCL files were processed running the *demux* command. Cell Ranger ATAC 2.0.0 was used to process Chromium Single Cell ATAC-seq data (10X Genomics), running *mkfastq*, and *cellranger-atac count* with reads aligned to mm10 reference genome (cellranger-arc-mm10-2020-A-2.0.0). Then, the peak/cell matrix was imported into Signac version 1.6.0 (https://satijalab.org/signac). Peaks were called using the Signac function *CallPeaks()*, which wraps the MACS2 function to call peaks. Doublet information was generated in ArchR (Granja et al. 2021) and was then exported and added to the Signac object.

Cells were retained based on the following QC metrics: peak_region_fragments > 0.02 & peak_region_fragments < 0.98, based on total sample percentages per term; pct_reads_in_peaks > 0.02; blacklist_ratio < 0.98; nucleosome_signal < 0.98; TSS.enrichment > 0.02. Common peaks were generated between biological replicates per timepoint / genotype and biological replicates merged within genotype and then between genotypes. After quality control and filtering, datasets from each of the three timepoints were analyzed as follows. Gene activities for each gene in each cell were calculated using the *GeneActivity*() function by summing the peak counts in the gene body + 2 kb upstream. Data were then normalized using term frequency inverse document frequency (TF-IDF) normalization (*RunTFIDF*), followed by dimensionality reduction using

Singular Value Decomposition (*RunSVD*). K-nearest neighbors were calculated using *FindNeighbors(reduction="lsi", dims=2:30)*. Finally, cell clusters were identified by a shared nearest neighbor (*SNN*) modularity optimization-based clustering algorithm *FindClusters(algorithm=3, resolution=0.6)*. To correct for batch differences, harmony was called using *RunHarmony(reduction = 'lsi', assay.use = 'peaks', group.by.vars = c('batch', 'orig.ident')*. UMAP was generated using *RunUMAP* function with *reduction="harmony*" and *dims=2:30*.

snRNA-seq and snATAC-seq data integration and cell-type annotations

Data from snATAC-seq were labeled based on cell labels in the corresponding snRNA-seq experiments described above at corresponding ages using Signac data integration. Briefly, cross-modality integration and label transfer was done using *FindTransferAnchors(reduction='cca')* and *TransferData(weight.reduction='lsi')* functions. Shared correlation patterns in the gene activity matrix and snRNA-seq datasets were used to match biological cell types across the two modalities. This analysis returned a classification of cell-type prediction scores for each cell. Cells were assigned the identity linked to their highest prediction score and poorly scored cells received a label of non-specific ("*Nonspec*").

Cell type re-clustering

Cells annotated as part of the excitatory neuronal lineage were subset as in the snRNA-seq dataset and processed as described above for re-clustering and cell-type annotation. This subset was used as the working excitatory neuronal lineage analysis for downstream analyses, with cell types annotated under "*CellType*" and "*CellType_coll*". This comprised 163,290 peaks and 9,111 nuclei at E13.5 from n=3/genotype; 165,957 peaks and 5,076 nuclei at E16.5 from n=2/genotype; and 188,540 peaks and 12,620 nuclei at P0.5 from n=3/genotype.

Differentially accessible regions

Differentially accessible regions (DARs) were calculated using the Signac package *FindMarkers* with *test.use = "LR"* and *latent.vars = c('peak_region_fragments','batch')* and compared for collapsed cell types "*CellType_coll*" of the excitatory neuronal lineage between genotypes per cell

type per timepoint. Peaks for X and Y chromosomes were removed. DARs were significant if the corrected p value was less than 0.05 (adj. *p-val < 0.05*) and |logFC| > 0.25).

Motif enrichment

To determine overrepresented motifs in a set of differentially accessible peaks, we followed the Signac motif analysis vignette. First, motif information was added to the each timepoint dataset using *AddMotifs*(). Motif information from mm10 was added from the JASPAR2020 database in Signac. We then calculated differential accessibility of each collapsed cell type between genotypes per timepoint, using *FindMarkers*() function with *test.use = 'LR'* and *latent.vars = ('nCount_peaks', 'batch')* in Signac. For each cell collapsed type, differential distal elements linked to genes with differential gene activity were used for motif enrichment analysis. Peaks were significant with p_val_adj <0.005 and |avg_log2FC| < 0.25. To find overrepresented motifs, we scanned a given set of differentially accessible peaks for all the DNA-binding motifs JASPAR2020_CORE_vertebrates_non-redundant databases (https://jaspar2020.genereg.net). Using *FindMotifs*(), we then computed the number of features containing the motif (*observed*) compared to the background. Background peaks were randomly sampled from all snATAC–seq peaks and matched for GC content using *MatchRegionStats()* in Signac. Enriched FOX motifs were tested for within overlapping DEG/DARs using FIMO from the MEME suite (Grant et al. 2011).

Data Availability

All acquired data are available upon request to the lead contact.

Code availability: https://github.com/anabrains/neocortex_genomics_Foxp1

Acknowledgments

Our sincerest thanks to Rachael Vollmer and Drs. Yuxiang Liu and Willy R. Vasquez for providing feedback on the manuscript. We also thank Dr. Shin Yamazaki and the Neuroscience Microscopy Facility, supported by the UT Southwestern Neuroscience Department and the UTSW Peter O'Donnell, Jr. Brain Institute. G.K. is a Jon Heighten Scholar in Autism Research and Townsend Distinguished Chair in Research on Autism Spectrum Disorders at UT Southwestern. This study was funded by UTSW T32 Molecular Medicine Research Training Program (NIH T32GM109776), R01MH102603 diversity supplement, F31MH123140-01, and HHMI Gilliam Fellowship to A.O.; the NIH (R01MH102603, R01MH126481), and the Simons Foundation for Autism Research Award (573689) to G.K.

Author Contributions

A.O. and G.K. designed the study. G.K. edited the manuscript. A.O. and F.A. collected and processed tissue for snRNA-seq. M.H. maintained mouse lines and performed genotyping. A.O. performed mouse husbandry; collected, and processed tissue for immuno-histochemistry, snATAC-seq, performed bioinformatic analyses, and wrote the manuscript.

References

Anderson AG, Kulkarni A, Harper M, Konopka G. 2020. Single-Cell Analysis of Foxp1-Driven Mechanisms Essential for Striatal Development. *Cell Reports* **30**: 3051-3066.e3057.

Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data.

Angerer P, Haghverdi L, Büttner M, Theis FJ, Marr C, Buettner F. 2016. destiny: diffusion maps for large-scale single-cell data in R. *Bioinformatics* **32**: 1241-1243.

Araujo DJ, Anderson AG, Berto S, Runnels W, Harper M, Ammanuel S, Rieger MA, Huang HC, Rajkovich K, Loerwald KW et al. 2015. FoxP1 orchestration of ASD-relevant signaling pathways in the striatum. *Genes Dev* **29**: 2081-2096.

Araujo DJ, Toriumi K, Escamilla CO, Kulkarni A, Anderson AG, Harper M, Usui N, Ellegood J, Lerch JP, Birnbaum SG et al. 2017. Foxp1 in Forebrain Pyramidal Neurons Controls Gene Expression Required for Spatial Learning and Synaptic Plasticity. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **37**: 10917-10931.

Ayhan F, Douglas C, Lega BC, Konopka G. 2021a. Nuclei isolation from surgically resected human hippocampus. *STAR Protocols* **2**: 100844.

Ayhan F, Kulkarni A, Berto S, Sivaprakasam K, Douglas C, Lega BC, Konopka G. 2021b. Resolving cellular and molecular diversity along the hippocampal anterior-to-posterior axis in humans. *Neuron* **109**: 2091-2105.e2096.

Bacon C, Schneider M, Le Magueresse C, Froehlich H, Sticht C, Gluch C, Monyer H, Rappold GA. 2015. Brain-specific Foxp1 deletion impairs neuronal development and causes autistic-like behaviour. *Mol Psychiatry* **20**: 632-639.

Banerjee-Basu S, Packer A. 2010. SFARI Gene: an evolving database for the autism research community. *Disease Models & Mechanisms* **3**: 133-135.

Betancourt J, Katzman S, Chen B. 2014. Nuclear factor one B regulates neural stem cell differentiation and axonal projection of corticofugal neurons. *Journal of Comparative Neurology* **522**: 6-35.

Borrell V, Calegari F. 2014. Mechanisms of brain evolution: regulation of neural progenitor cell diversity and cell cycle length. *Neurosci Res* **86**: 14-24.

Braccioli L, Vervoort SJ, Adolfs Y, Heijnen CJ, Basak O, Pasterkamp RJ, Nijboer CH, Coffer PJ. 2017. FOXP1 Promotes Embryonic Neural Stem Cell Differentiation by Repressing Jagged1 Expression. *Stem Cell Reports* **9**: 1530-1545.

Bunt J, Osinski JM, Lim JW, Vidovic D, Ye Y, Zalucki O, O'Connor TR, Harris L, Gronostajski RM, Richards LJ et al. 2017. Combined allelic dosage of Nfia and Nfib regulates cortical development. *Brain Neurosci Adv* **1**: 2398212817739433.

Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. 2018. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology* **36**: 411-420.

Caglayan E, Liu Y, Konopka G. 2022. Neuronal ambient RNA contamination causes misinterpreted and masked cell types in brain single-nuclei datasets. *Neuron* **110**: 4043-4056.e4045.

Cardo LF, de la Fuente DC, Li M. 2023. Impaired neurogenesis and neural progenitor fate choice in a human stem cell model of SETBP1 disorder. *Molecular Autism* **14**: 8.

Carr CW, Moreno-De-Luca D, Parker C, Zimmerman HH, Ledbetter N, Martin CL, Dobyns WB, Abdul-Rahman OA. 2010. Chiari I malformation, delayed gross motor skills, severe speech delay, and epileptiform discharges in a child with FOXP1 haploinsufficiency. *Eur J Hum Genet* **18**: 1216-1220.

Chen G, Sima J, Jin M, Wang K-y, Xue X-j, Zheng W, Ding Y-q, Yuan X-b. 2008. Semaphorin-3A guides radial migration of cortical neurons during development. *Nature neuroscience* **11**: 36-44.

Chen L, Wang Y, Jiang W, Ni R, Wang Y, Ni S. 2019. CDK14 involvement in proliferation migration and invasion of esophageal cancer. *Ann Transl Med* **7**: 681.

Chokas AL, Trivedi CM, Lu MM, Tucker PW, Li S, Epstein JA, Morrisey EE. 2010. Foxp1/2/4-NuRD Interactions Regulate Gene Expression and Epithelial Injury Response in the Lung via Regulation of Interleukin-6*. *Journal of Biological Chemistry* **285**: 13304-13313.

de la Torre-Ubieta L, Won H, Stein JL, Geschwind DH. 2016. Advancing the understanding of autism disease mechanisms through genetics. *NATURE MEDICINE*.

Di Bella DJ, Habibi E, Stickels RR, Scalia G, Brown J, Yadollahpour P, Yang SM, Abbate C, Biancalani T, Macosko EZ et al. 2021. Molecular logic of cellular diversification in the mouse cerebral cortex. *Nature* **595**: 554-559.

Estruch SB, Graham SA, Quevedo M, Vino A, Dekkers DHW, Deriziotis P, Sollis E, Demmers J, Poot RA, Fisher SE. 2018. Proteomic analysis of FOXP proteins reveals interactions between cortical transcription factors associated with neurodevelopmental disorders. *Human molecular genetics* **27**: 1212-1227.

Feng X, Ippolito GC, Tian L, Wiehagen K, Oh S, Sambandam A, Willen J, Bunte RM, Maika SD, Harriss JV et al. 2010. Foxp1 is an essential transcriptional regulator for the generation of quiescent naive T cells during thymocyte development. *Blood* **115**: 510-518.

Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK, Slichter CK, Miller HW, McElrath MJ, Prlic M et al. 2015. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biology* **16**: 278.

Fleming SJ, Chaffin MD, Arduini A, Akkad AD, Banks E, Marioni JC, Philippakis AA, Ellinor PT, Babadi M. 2023. Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender. *Nature methods* **20**: 1323-1335.

Franco SJ, Gil-Sanz C, Martinez-Garay I, Espinosa A, Harkins-Perry SR, Ramos C, Müller U. 2012. Fate-Restricted Neural Progenitors in the Mammalian Cerebral Cortex. *Science* **337**: 746-749.

Frantz GD, McConnell SK. 1996. Restriction of Late Cerebral Cortical Progenitors to an Upper-Layer Fate. *Neuron* **17**: 55-61.

Fröhlich H, Rafiullah R, Schmitt N, Abele S, Rappold GA. 2017. Foxp1 expression is essential for sex-specific murine neonatal ultrasonic vocalization. *Human molecular genetics* **26**: 1511-1521.

Fu JM, Satterstrom FK, Peng M, Brand H, Collins RL, Dong S, Wamsley B, Klei L, Wang L, Hao SP et al. 2022. Rare coding variation provides insight into the genetic architecture and phenotypic context of autism. *Nature genetics* **54**: 1320-1331.

Gabut M, Samavarchi-Tehrani P, Wang X, Slobodeniuc V, O'Hanlon D, Sung H, Alvarez M, Talukder S, Pan Q, Mazzoni EO et al. 2011. An Alternative Splicing Switch Regulates Embryonic Stem Cell Pluripotency and Reprogramming. *Cell* **147**: 132-146.

Gao P, Postiglione MP, Krieger TG, Hernandez L, Wang C, Han Z, Streicher C, Papusheva E, Insolera R, Chugh K et al. 2014. Deterministic progenitor behavior and unitary production of neurons in the neocortex. *Cell* **159**: 775-788.

Gogolla N, Leblanc JJ, Quast KB, Südhof TC, Fagiolini M, Hensch TK. 2009. Common circuit defect of excitatory-inhibitory balance in mouse models of autism. *J Neurodev Disord* **1**: 172-181.

Gorski JA, Talley T, Qiu M, Puelles L, Rubenstein JLR, Jones KR. 2002. Cortical Excitatory Neurons and Glia, But Not GABAergic Neurons, Are Produced in the Emx1-Expressing Lineage. *The Journal of Neuroscience* **22**: 6309-6314.

Götz M, Huttner WB. 2005. The cell biology of neurogenesis. *Nature Reviews Molecular Cell Biology* **6**: 777-788.

Govindan S, Oberst P, Jabaudon D. 2018. In vivo pulse labeling of isochronic cohorts of cells in the central nervous system using FlashTag. *Nature Protocols* **13**: 2297-2311.

Granja JM, Corces MR, Pierce SE, Bagdatli ST, Choudhry H, Chang HY, Greenleaf WJ. 2021. ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nature genetics* **53**: 403-411.

Grant CE, Bailey TL, Noble WS. 2011. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**: 1017-1018.

Habib N, Avraham-Davidi I, Basu A, Burks T, Shekhar K, Hofree M, Choudhury SR, Aguet F, Gelfand E, Ardlie K et al. 2017. Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nature methods* **14**: 955-958.

Hamdan FF, Daoud H, Rochefort D, Piton A, Gauthier J, Langlois M, Foomani G, Dobrzeniecka S, Krebs MO, Joober R et al. 2010. De novo mutations in FOXP1 in cases with intellectual disability, autism, and language impairment. *The American Journal of Human Genetics* **87**: 671-678.

Hanashima C, Shen L, Li SC, Lai E. 2002. Brain factor-1 controls the proliferation and differentiation of neocortical progenitor cells through independent mechanisms. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **22**: 6526-6536.

Heng YHE, Barry G, Richards LJ, Piper M. 2012. Nuclear Factor I Genes Regulate Neuronal Migration. *Neurosignals* **20**: 159-167.

Hettige NC, Peng H, Wu H, Zhang X, Yerko V, Zhang Y, Jefri M, Soubannier V, Maussion G, Alsuwaidi S et al. 2022. FOXG1 dose tunes cell proliferation dynamics in human forebrain progenitor cells. *Stem Cell Reports* **17**: 475-488.

Hickey SL, Berto S, Konopka G. 2019. Chromatin Decondensation by FOXP2 Promotes Human Neuron Maturation and Expression of Neurodevelopmental Disease Genes. *Cell Reports* **27**: 1699-1711.e1699.

Hisaoka T, Nakamura Y, Senba E, Morikawa Y. 2010. The forkhead transcription factors, Foxp1 and Foxp2, identify different subpopulations of projection neurons in the mouse cerebral cortex. *Neuroscience* **166**: 551-563.

Jepsen K, Gleiberman AS, Shi C, Simon DI, Rosenfeld MG. 2008. Cooperative regulation in development by SMRT and FOXP1. *Genes Dev* **22**: 740-745.

Katayama Y, Nishiyama M, Shoji H, Ohkawa Y, Kawamura A, Sato T, Suyama M, Takumi T, Miyakawa T, Nakayama KI. 2016. CHD8 haploinsufficiency results in autistic-like phenotypes in mice. *Nature* **537**: 675-679.

Khandelwal N, Cavalier S, Rybalchenko V, Kulkarni A, Anderson AG, Konopka G, Gibson JR. 2021. FOXP1 negatively regulates intrinsic excitability in D2 striatal projection neurons by promoting inwardly rectifying and leak potassium currents. *Mol Psychiatry* **26**: 1761-1774.

Klingler E, Francis F, Jabaudon D, Cappello S. 2021. Mapping the molecular and cellular complexity of cortical malformations. *Science* **371**.

Knauss JL, Miao N, Kim SN, Nie Y, Shi Y, Wu T, Pinto HB, Donohoe ME, Sun T. 2018. Long noncoding RNA Sox2ot and transcription factor YY1 co-regulate the differentiation of cortical neural progenitors by repressing Sox2. *Cell Death Dis* **9**: 799.

Kowalczyk T, Pontious A, Englund C, Daza RAM, Bedogni F, Hodge R, Attardo A, Bell C, Huttner WB, Hevner RF. 2009. Intermediate Neuronal Progenitors (Basal Progenitors) Produce Pyramidal–Projection Neurons for All Layers of Cerebral Cortex. *Cerebral Cortex* **19**: 2439-2450.

Kriegstein AR, Götz M. 2003. Radial glia diversity: a matter of cell fate. *Glia* **43**: 37-43.

Li S, Weidenfeld J, Morrisey EE. 2004. Transcriptional and DNA binding activity of the Foxp1/2/4 family is modulated by heterotypic and homotypic protein interactions. *Mol Cell Biol* **24**: 809-822.

Li X, Xiao J, Fröhlich H, Tu X, Li L, Xu Y, Cao H, Qu J, Rappold GA, Chen J-G. 2015. Foxp1 regulates cortical radial migration and neuronal morphogenesis in developing cerebral cortex. *PloS one* **10**: e0127671-e0127671.

Li Y, Wang R, Qiao N, Peng G, Zhang K, Tang K, Han JJ, Jing N. 2017. Transcriptome analysis reveals determinant stages controlling human embryonic stem cell commitment to neuronal cells. *The Journal of biological chemistry* **292**: 19590-19604.

Limoni G, Murthy S, Jabaudon D, Dayer A, Niquille M. 2021. PlexinA4-Semaphorin3A-mediated crosstalk between main cortical interneuron classes is required for superficial interneuron lamination. *Cell Rep* **34**: 108644.

Limoni G, Niquille M. 2021. Semaphorins and Plexins in central nervous system patterning: the key to it all? *Curr Opin Neurobiol* **66**: 224-232.

Loo L, Simon JM, Xing L, McCoy ES, Niehaus JK, Guo J, Anton ES, Zylka MJ. 2019. Single-cell transcriptomic analysis of mouse neocortical development. *Nat Commun* **10**: 134.

Lozano R, Gbekie C, Siper PM, Srivastava S, Saland JM, Sethuram S, Tang L, Drapeau E, Frank Y, Buxbaum JD et al. 2021. FOXP1 syndrome: a review of the literature and practice parameters for medical assessment and monitoring. *J Neurodev Disord* **13**: 18.

Malatesta P, Hartfuss E, Götz M. 2000. Isolation of radial glial cells by fluorescent-activated cell sorting reveals a neuronal lineage. *Development* **127**: 5253-5263.

Martynoga B, Mateo JL, Zhou B, Andersen J, Achimastou A, Urbán N, van den Berg D, Georgopoulou D, Hadjur S, Wittbrodt J et al. 2013. Epigenomic enhancer annotation reveals a key role for NFIX in neural stem cell quiescence. *Genes Dev* **27**: 1769-1786.

Martynoga B, Morrison H, Price DJ, Mason JO. 2005. Foxg1 is required for specification of ventral telencephalon and region-specific regulation of dorsal telencephalic precursor proliferation and apoptosis. *Dev Biol* **283**: 113-127.

Meerschaut I, Rochefort D, Revençu N, Pètre J, Corsello C, Rouleau GA, Hamdan FF, Michaud JL, Morton J, Radley J et al. 2017. FOXP1-related intellectual disability syndrome: a recognisable entity. *Journal of Medical Genetics* **54**: 613.

Mihalas AB, Elsen GE, Bedogni F, Daza RAM, Ramos-Laguna KA, Arnold SJ, Hevner RF. 2016. Intermediate Progenitor Cohorts Differentially Generate Cortical Layers and Require Tbr2 for Timely Acquisition of Neuronal Subtype Identity. *Cell Rep* **16**: 92-105.

Molyneaux BJ, Arlotta P, Menezes JRL, Macklis JD. 2007. Neuronal subtype specification in the cerebral cortex. *Nature Reviews Neuroscience* **8**: 427-437.

Morgan A, Braden R, Wong MMK, Colin E, Amor D, Liégeois F, Srivastava S, Vogel A, Bizaoui V, Ranguin K et al. 2021. Speech and language deficits are central to SETBP1 haploinsufficiency disorder. *Eur J Hum Genet* **29**: 1216-1225.

O'Roak BJ, Deriziotis P, Lee C, Vives L, Schwartz JJ, Girirajan S, Karakoc E, MacKenzie AP, Ng SB, Baker C et al. 2011. Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nature genetics* **43**: 585-589.

Paap RH, Oosterbroek S, Wagemans C, von Oerthel L, Schellevis RD, Vastenhouw-van der Linden AJA, Groot Koerkamp MJA, Hoekman MFM, Smidt MP. 2016. FoxO6 affects Plxna4-mediated neuronal migration during mouse cortical development. *Proc Natl Acad Sci U S A* **113**: E7087-e7096.

Pariani MJ, Spencer A, Graham JM, Jr., Rimoin DL. 2009. A 785kb deletion of 3p14.1p13, including the FOXP1 gene, associated with speech delay, contractures, hypertonia and blepharophimosis. *Eur J Med Genet* **52**: 123-127.

Park SHE, Kulkarni A, Konopka G. 2023. FOXP1 orchestrates neurogenesis in human cortical basal radial glial cells. *PLoS Biol* **21**: e3001852.

Park SHE, Ortiz AK, Konopka G. 2022. Corticogenesis across species at single-cell resolution. *Dev Neurobiol* **82**: 517-532.

Paul V, Tonchev AB, Henningfeld KA, Pavlakis E, Rust B, Pieler T, Stoykova A. 2014. Scratch2 modulates neurogenesis and cell migration through antagonism of bHLH proteins in the developing neocortex. *Cereb Cortex* **24**: 754-772.

Paulsen B, Velasco S, Kedaigle AJ, Pigoni M, Quadrato G, Deo AJ, Adiconis X, Uzquiano A, Sartore R, Yang SM et al. 2022. Autism genes converge on asynchronous development of shared neuron classes. *Nature* **602**: 268-273.

Pearson CA, Moore DM, Tucker HO, Dekker JD, Hu H, Miquelajáuregui A, Novitch BG. 2020. Foxp1 Regulates Neural Stem Cell Self-Renewal and Bias Toward Deep Layer Cortical Fates. *Cell Reports* **30**: 1964-1981.e1963.

Quesnel-Vallières M, Weatheritt RJ, Cordes SP, Blencowe BJ. 2019. Autism spectrum disorder: insights into convergent mechanisms from transcriptomics. *Nature Reviews Genetics* **20**: 51-63.

Roberts JZ, Crawford N, Longley DB. 2022. The role of Ubiquitination in Apoptosis and Necroptosis. *Cell Death Differ* **29**: 272-284.

Rubenstein JL, Merzenich MM. 2003. Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes Brain Behav* **2**: 255-267.

Satterstrom FK, Kosmicki JA, Wang J, Breen MS, De Rubeis S, An JY, Peng M, Collins R, Grove J, Klei L et al. 2020. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell* **180**: 568-584.e523.

Shen Q, Wang Y, Dimos JT, Fasano CA, Phoenix TN, Lemischka IR, Ivanova NB, Stifani S, Morrisey EE, Temple S. 2006. The timing of cortical neurogenesis is encoded within lineages of individual progenitor cells. *Nature neuroscience* **9**: 743-751.

Simon R, Wiegreffe C, Britsch S. 2020. Bcl11 Transcription Factors Regulate Cortical Development and Function. *Front Mol Neurosci* **13**: 51.

Siper PM, De Rubeis S, Trelles MDP, Durkin A, Di Marino D, Muratet F, Frank Y, Lozano R, Eichler EE, Kelly M et al. 2017. Prospective investigation of FOXP1 syndrome. *Mol Autism* **8**: 57.

Sorokina O, Sorokin A, Armstrong JD. 2022. Synaptome.db: a bioconductor package for synaptic proteomics data. *Bioinformatics Advances* **2**: vbac086.

Telley L, Agirman G, Prados J, Amberg N, Fièvre S, Oberst P, Bartolini G, Vitali I, Cadilhac C, Hippenmeyer S et al. 2019. Temporal patterning of apical progenitors and their daughter neurons in the developing neocortex. *Science* **364**: eaav2522.

Usui N, Araujo DJ, Kulkarni A, Co M, Ellegood J, Harper M, Toriumi K, Lerch JP, Konopka G. 2017. Foxp1 regulation of neonatal vocalizations via cortical development. *Genes & development* **31**: 2039-2055.

Vasistha NA, García-Moreno F, Arora S, Cheung AF, Arnold SJ, Robertson EJ, Molnár Z. 2015. Cortical and Clonal Contribution of Tbr2 Expressing Progenitors in the Developing Mouse Brain. *Cereb Cortex* **25**: 3290-3302.

Vitali I, Fièvre S, Telley L, Oberst P, Bariselli S, Frangeul L, Baumann N, McMahon JJ, Klingler E, Bocchi R et al. 2018. Progenitor Hyperpolarization Regulates the Sequential Generation of Neuronal Subtypes in the Developing Neocortex. *Cell* **174**: 1264-1276.e1215.

Wang J, Fröhlich H, Torres FB, Silva RL, Poschet G, Agarwal A, Rappold GA. 2022a. Mitochondrial dysfunction and oxidative stress contribute to cognitive and motor impairment in FOXP1 syndrome. *Proc Natl Acad Sci U S A* **119**.

Wang J, Zhai HR, Ma SF, Shi HZ, Zhang WJ, Yun Q, Liu WJ, Liu ZZ, Zhang WN. 2022b. FOXG1 Contributes Adult Hippocampal Neurogenesis in Mice. *Int J Mol Sci* **23**.

Wang S, Tan N, Zhu X, Yao M, Wang Y, Zhang X, Xu Z. 2018. Sh3rf2 Haploinsufficiency Leads to Unilateral Neuronal Development Deficits and Autistic-Like Behaviors in Mice. *Cell Rep* **25**: 2963-2971.e2966.

Wilhelm M, Kukekov NV, Schmit TL, Biagas KV, Sproul AA, Gire S, Maes ME, Xu Z, Greene LA. 2012. Sh3rf2/POSHER protein promotes cell survival by ring-mediated proteasomal degradation of the c-Jun N-terminal kinase scaffold POSH (Plenty of SH3s) protein. *The Journal of biological chemistry* **287**: 2247-2256.

Wong FK, Fei JF, Mora-Bermúdez F, Taverna E, Haffner C, Fu J, Anastassiadis K, Stewart AF, Huttner WB. 2015. Sustained Pax6 Expression Generates Primate-like Basal Radial Glia in Developing Mouse Neocortex. *PLoS Biol* **13**: e1002217.

Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L et al. 2021. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)* **2**: 100141.

Yoon KJ, Ringeling FR, Vissers C, Jacob F, Pokrass M, Jimenez-Cyrus D, Su Y, Kim NS, Zhu Y, Zheng L et al. 2017. Temporal Control of Mammalian Cortical Neurogenesis by m(6)A Methylation. *Cell* **171**: 877-889.e817.

Zhang Y, Li S, Yuan L, Tian Y, Weidenfeld J, Yang J, Liu F, Chokas AL, Morrisey EE. 2010. Foxp1 coordinates cardiomyocyte proliferation through both cell-autonomous and nonautonomous mechanisms. *Genes Dev* **24**: 1746-1757.

Zheng GXY, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, Ziraldo SB, Wheeler TD, McDermott GP, Zhu J et al. 2017. Massively parallel digital transcriptional profiling of single cells. in *Nat Commun*, p. 14049.

Figures

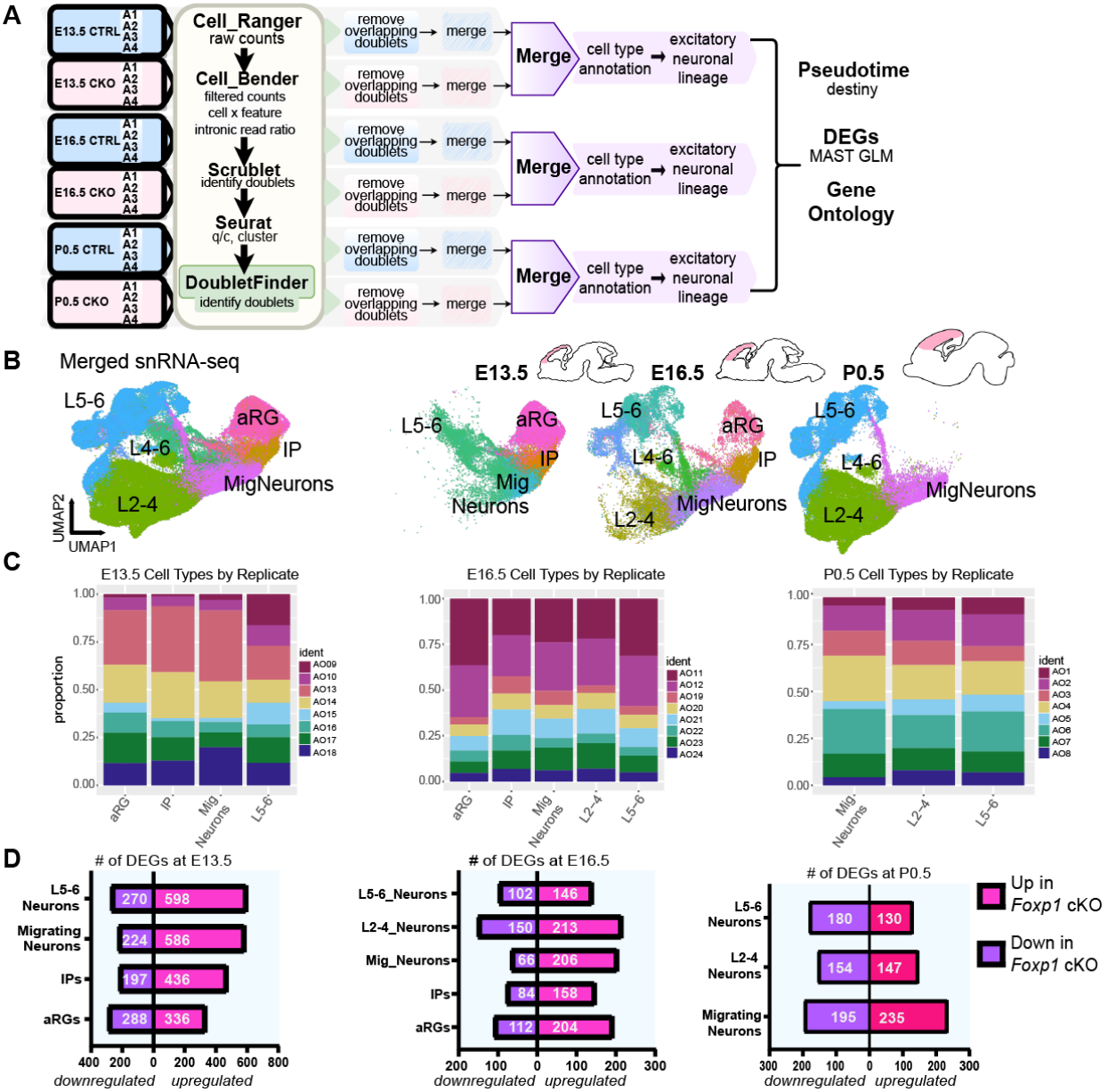

Figure 1: Excitatory neuronal lineage snRNA-seq of *Foxp1* cKO and control cortices shows disruptions in gene expression. **A)** Visual pipeline for snRNA-seq data processing of n=4 replicates per genotype per timepoint (E13.5, E16.5, P0.5) of neocortex dissections. **B)** Processed snRNA-seq data from all timepoints demonstrates expected cell types and cell distributions across neocortical development. Total cells per timepoint are E13.5: 31,543; E16.5: 26,444; P0.5: 50,672. **C)** Stacked bar plots show representation from replicates across broad cell-type clusters. **D)** Number of DEGs per broad cell type by timepoint. DEGs were defined as significant with adj. p-val < 0.05 and average |logFC| > 0.15.

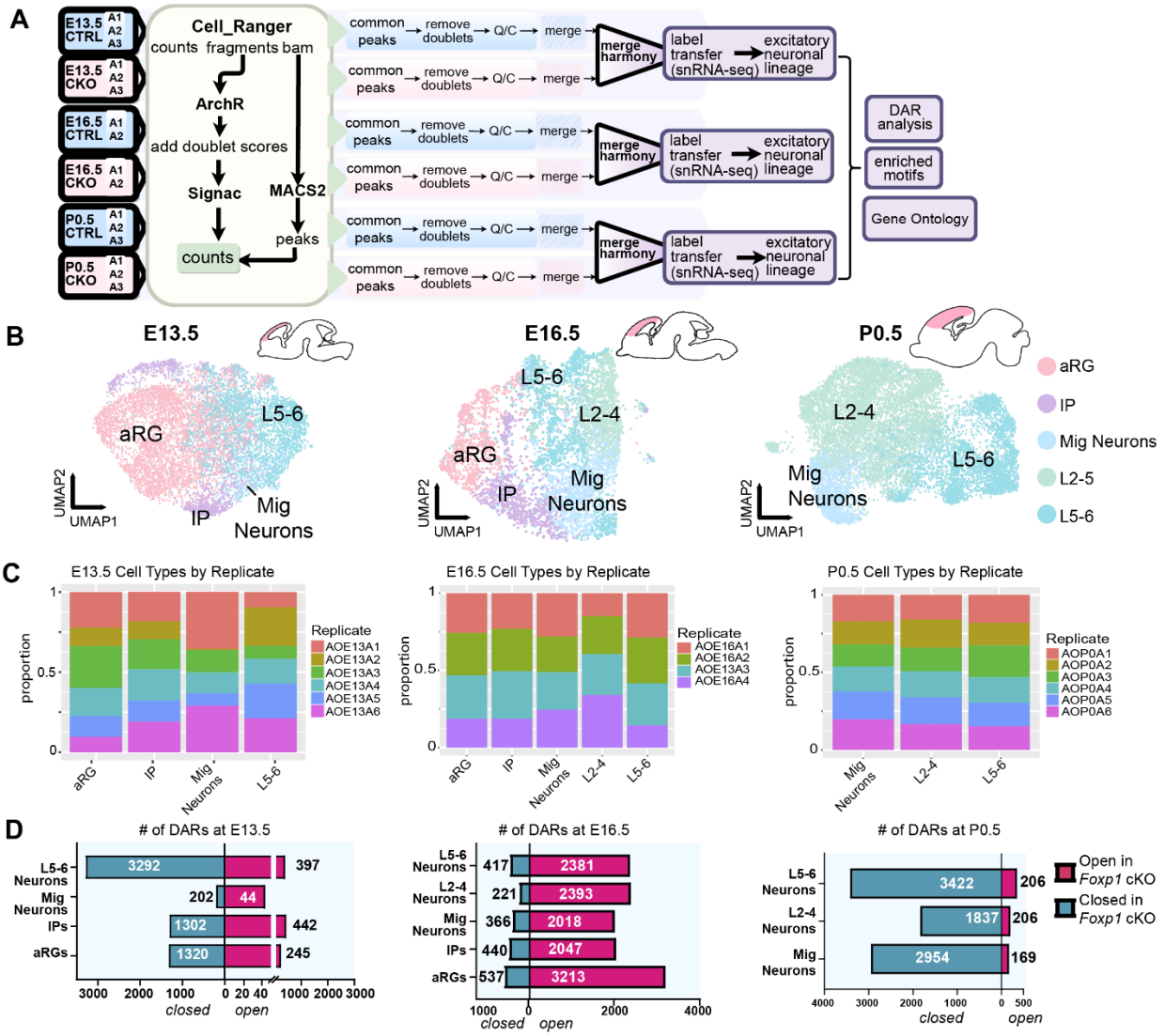

Figure 1: Excitatory neuronal lineage snATAC-seq of *Foxp1* cKO and control cortices shows disruptions in chromatin accessibility. **A)** Visual pipeline for snATAC-seq data processing per timepoint (E13.5, E16.5, P0.5) of whole cortex dissections. **B**) Processed snATAC-seq data demonstrates expected cell types and cell distributions per time-point across neocortical development comparable to snRNA-seq data. 9,111 nuclei at E13.5 from n=3/genotype; 5,076 nuclei at E16.5 from n=2/genotype; and 12,620 nuclei at P0.5 from n=3/genotype. **C)** Stacked bar plots show representation from individual replicates across broad cell-type clusters. **D)** Number of differentially accessible regions (DARs) per broad cell-type timepoint highlights shifting role of *Foxp1* regulation. DARs were defined as significant with adj. p-val < 0.05 and avg |log2FC| > 0.25.

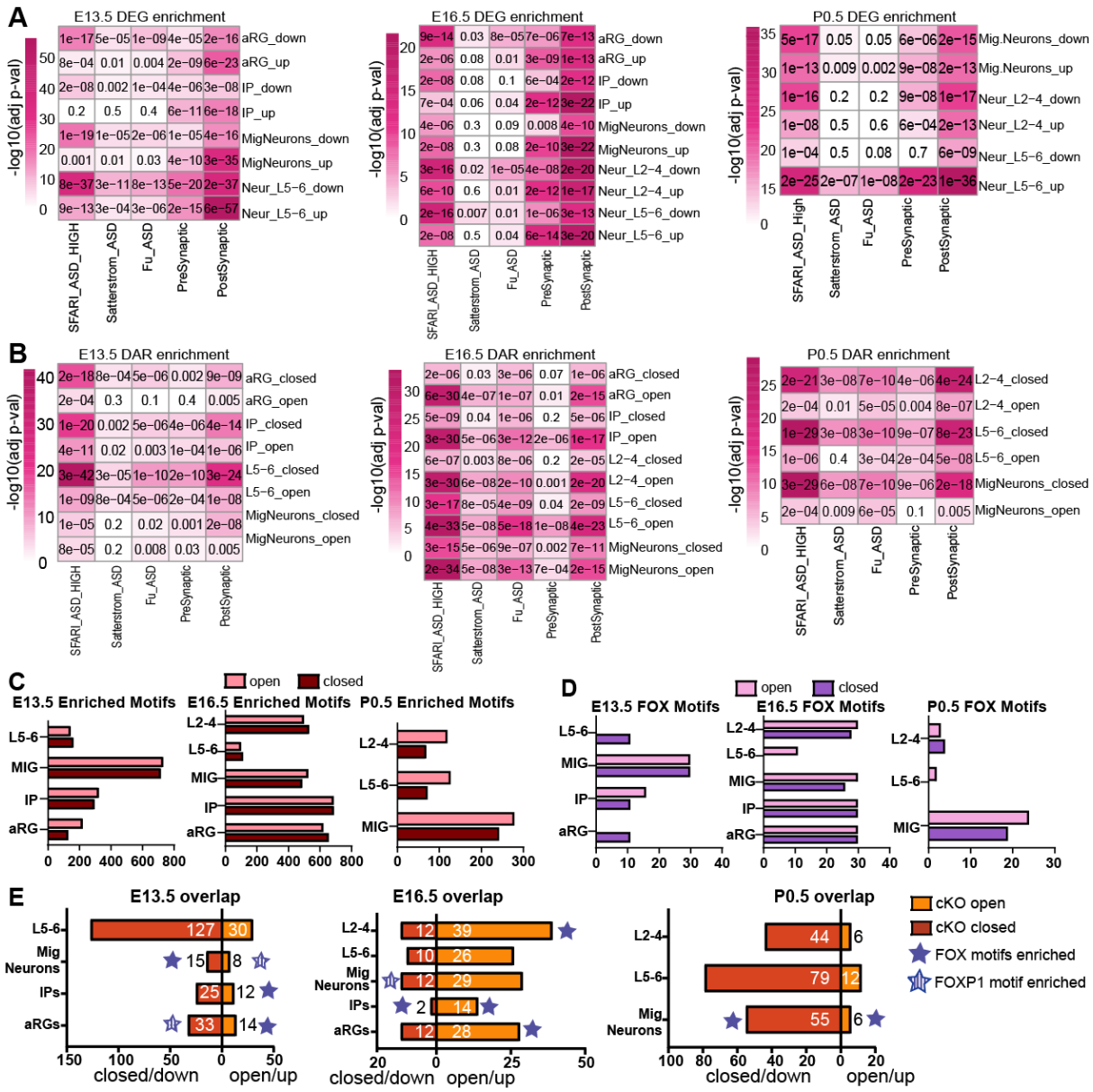

Figure 3: DEGs and DARs are associated with ASD-relevant and synaptic genes and overlapping DEGs/DARs are enriched for FOX motifs. A, B) Hypergeometric enrichment output of **(A)** DEGs or **(B)** DARs with high-confidence SFARI-ASD lists, ASD genes from Satterstrom et. al. or Fu et al., and pre- or post-synaptic genes from Synaptome.db. **A)** In synaptic gene datasets, DEGs have a higher enrichment with post-synaptic genes. **B)** Enrichment is highest for closed DARs and SFARI-ASD genes and synaptic genes at E13.5 and P0.5, with the reverse at E16.5. Hypergeometric enrichment was significant with adj. p-val < 0.05 and -$\log_{10}$(FDR)>1.31. The number in the boxes indicates the adj. p-val. **C)** Number of enriched motifs in open or closed differentially accessible peaks. Defined as significant with adj. p-val < 0.005 and avg |log2FC| > 0.25. **D)** Number of enriched FOX motifs. **E)** Number of overlapping downregulated DEGs / closed DARs or upregulated DEGs / open DARs. Star symbol indicates the presence of enriched FOX motifs in this specific cell type. The striped star specifies enriched FOXP1 motif.

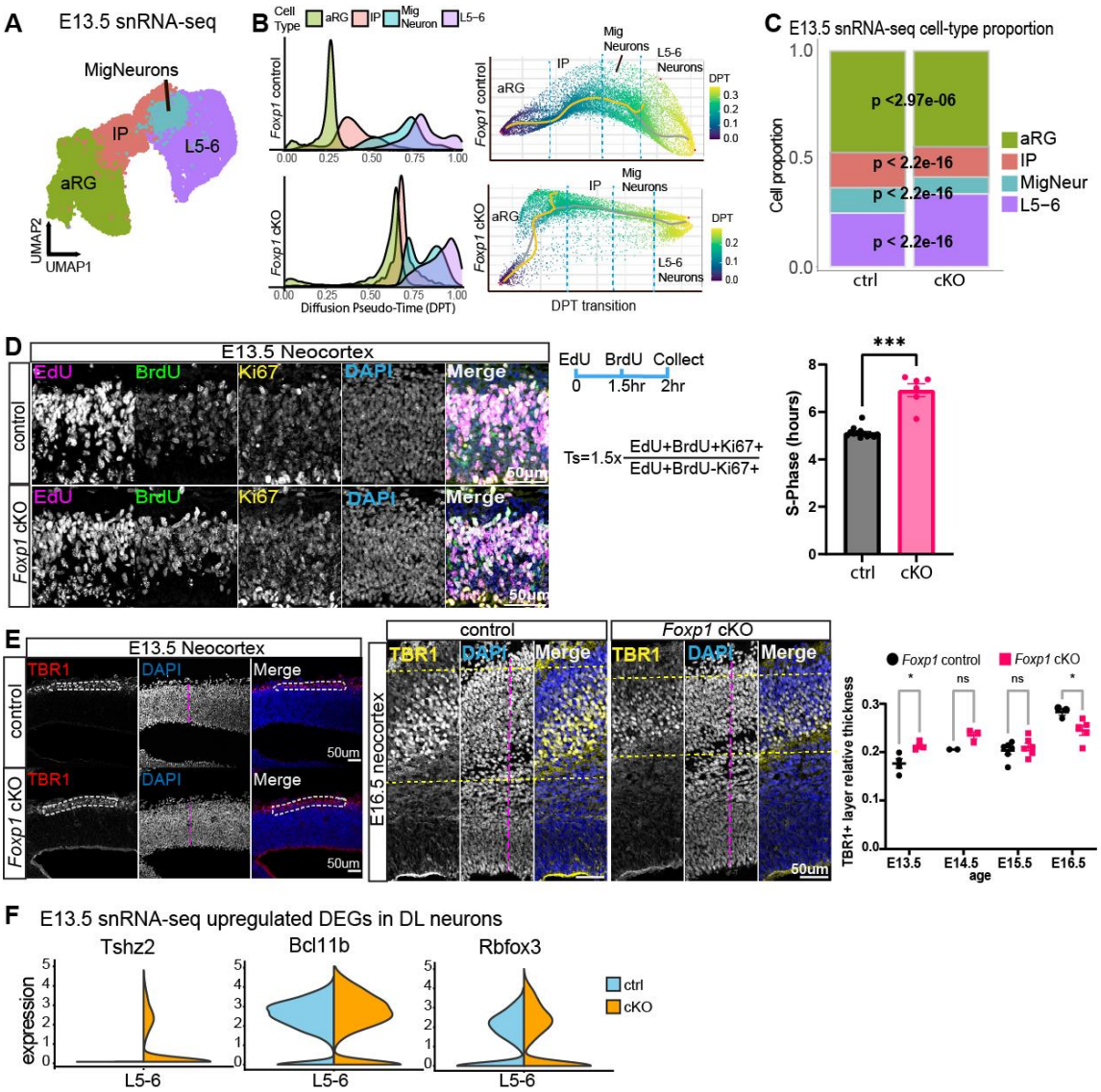

Figure 4: *Foxp1* cKO results in neurons with accelerated pseudo-age and precocious neurogenesis. A) UMAP plot of E13.5 snRNA-seq data with excitatory neuronal lineage broad cell-type annotation. **B)** *(Left)* Pseudotime plot of cell-type density distribution vs pseudotime. Path lengths between control are significantly different p < 2e-16 with Mann-Whitey test with continuity correction. *(Right)* Diffusion pseudotime axes showing pseudo-age and cell types. **C)** Stacked bar plot of cell-type proportion by genotype of E13.5 snRNA-seq data. p-values of two-proportions z-test with Yates correction. *Foxp1* cKOs have significantly reduced proportion of apical radial glia (aRG), intermediate progenitors (IPs), migrating neurons, and significantly higher proportion of L5-6 neurons. **D)** Estimation of S-Phase length shows an elongated S-Phase length in *Foxp1* cKOs at E13.5. Mean±SEM. **E)** IHC images at 20x of TBR1+ layer thickness relative to DAPI+ thickness at E13.5 (*left*) and E16.5 (*middle*) and quantification of TBR1+ relative layer thickness from E13.5-E16.5 (*right*). Mean±SEM. **F)** DL genes *Tshz2* and *Bcl11b*/*Ctip2* and general neuronal marker *Rbfox3*/*Neun* are upregulated in DL neurons of cKOs. Scale bars = 50μm.

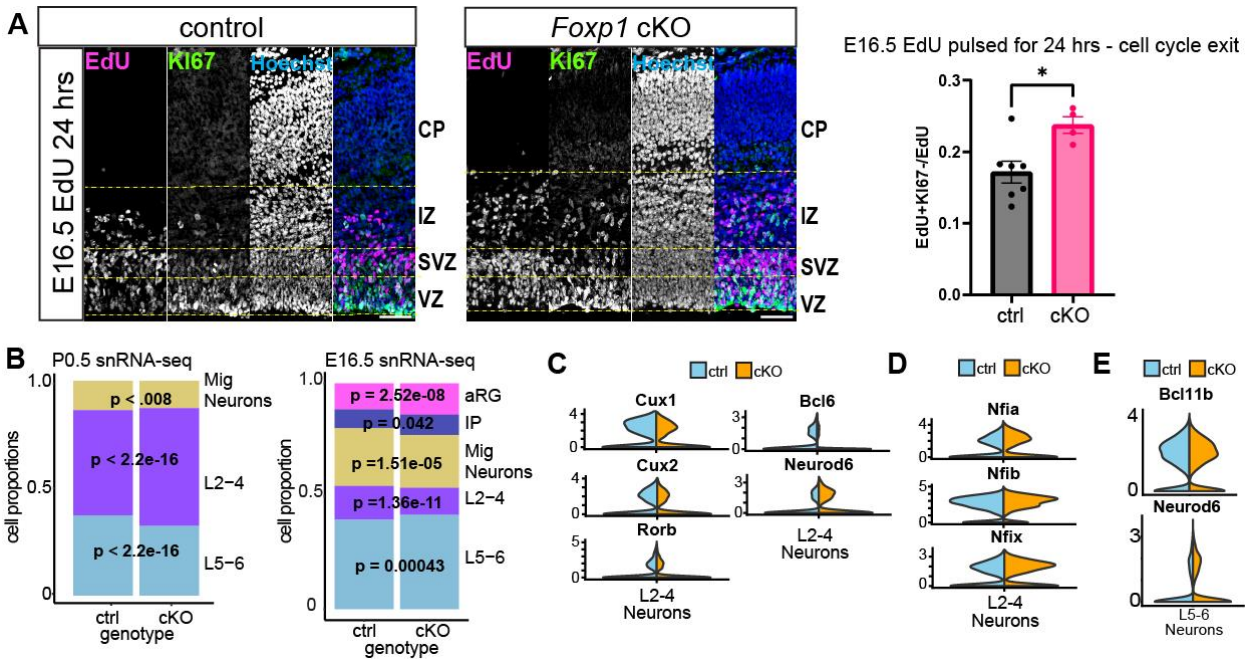

Figure 5: Accelerated cell cycle exit at E16.5 in *Foxp1* cKOs and altered cortical layer marker gene expression. A) *(Left*) IHC images show cell cycle exit assay of 24hr EdU pulse from E15.5 to E16.5. KI67: proliferative marker; Hoechst used as DNA marker. *(Right)* EdU+KI67- cells over total EdU cells; Cells that are EdU+ were cycling at the time of EdU injection. KI67- cells are no longer proliferating. EdU+KI67- cells indicate cells that have exited the cell cycle. *Foxp1* cKOs have higher EdU+Ki67-/EdU cells. Mean±SEM. VZ: ventricular zone. SVZ: sub-ventricular zone, IZ: intermediate zone, CP: cortical plate. **B)** (*left*) Cell-type proportions at P0.5 for snRNA-seq data by broad cell type per genotype. Faster cell-cycle exit during the typical generation of UL neurons (as in panel A) may underlie the postnatal phenotype of relatively thicker L2-4, and relatively thinner L5-6. (*right*) However, these differences are not detected at E16.5 in the snRNA-seq data likely since neuronal identity is still being refined. **C)** Altered expression of genes in E16.5 UL neurons: UL TF expression *Cux1*, *Cux2*, *Rorb*, *Bcl6* are downregulated in cKOs. Neuronal marker *Neurod6* is upregulated in cKOs. **D)** Enrichment of genes in E16.5 DL neuron clusters: *Bcl11b* is downregulated in cKOs and *Neurod6* is upregulated. **E)** At E16.5, migration relevant genes *Nfia/b/x* are upregulated in cKOs. Scale bar = 50µm.

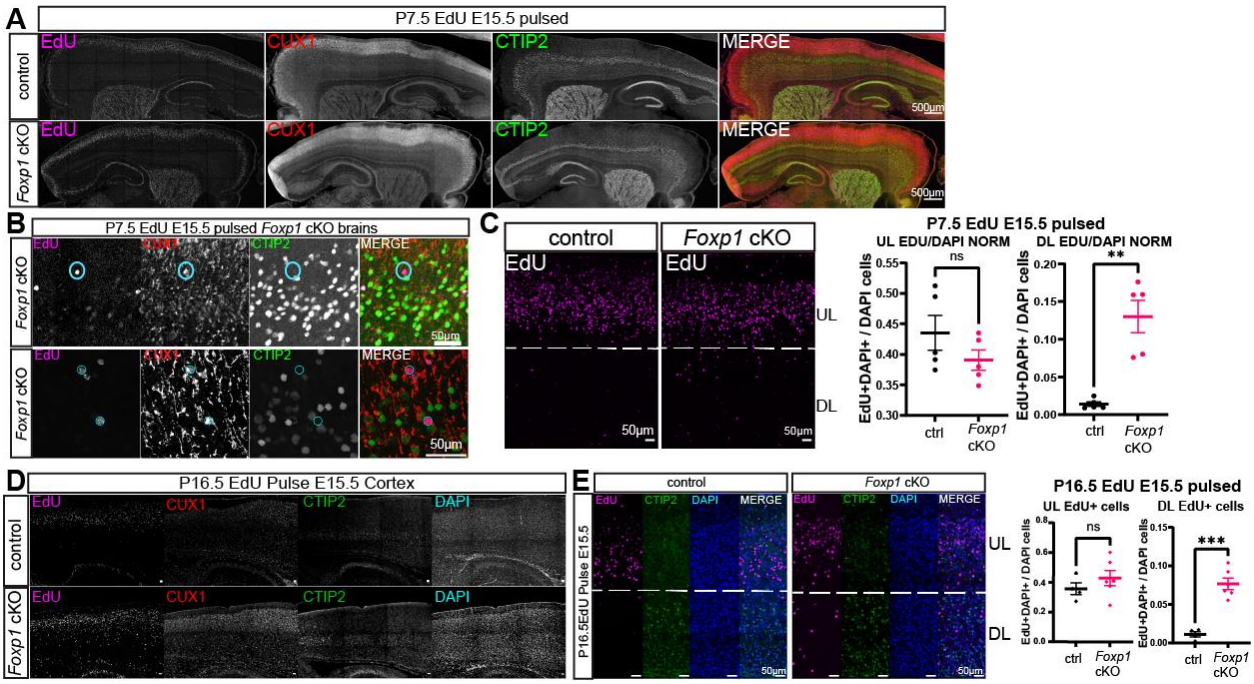

Figure 6: Selective neuronal lineage migration deficits but no change in timing of neurogenesis in *Foxp1* cKO animals. A-B) EdU birth-dating experiment shows that neurons generated from progenitors dividing at E15.5 remain in the DLs in *Foxp1* cKOs at higher levels than in controls throughout the cortex. Some of these cells are EdU+CUX1+ indicating they have an UL cell type identity and are CTIP2-. **B)** *Foxp1* cKO cortices *(top)* is at 20x and (*bottom*) at 63x showing EdU+CUX1+CTIP2- cells in L5-6 region. **C)** P7.5 brains stained for EdU that was pulsed at E15.5 have significantly more cells in the DLs in *Foxp1* cKOs. No significant difference between EdU+ cells in the ULs between genotypes. **D)** P16.5 brains from E15.5 EdU pulsed animals with CUX1 and CTIP2 to demarcate cortical layer boundaries at 10x resolution. **E)** P16.5 brains with EdU cell quantification in ULs or DLs shows a significant number of EdU+ cells located in the DLs in cKOs but not controls. No significant change in the number of UL EdU+ cells. EdU+ cells were CTIP2- despite being in L5-6, suggesting appropriate timing of neurogenesis at E15.5 for UL cells and a likely selective migration deficit that remains at P16.5. Mean±SEM shown.

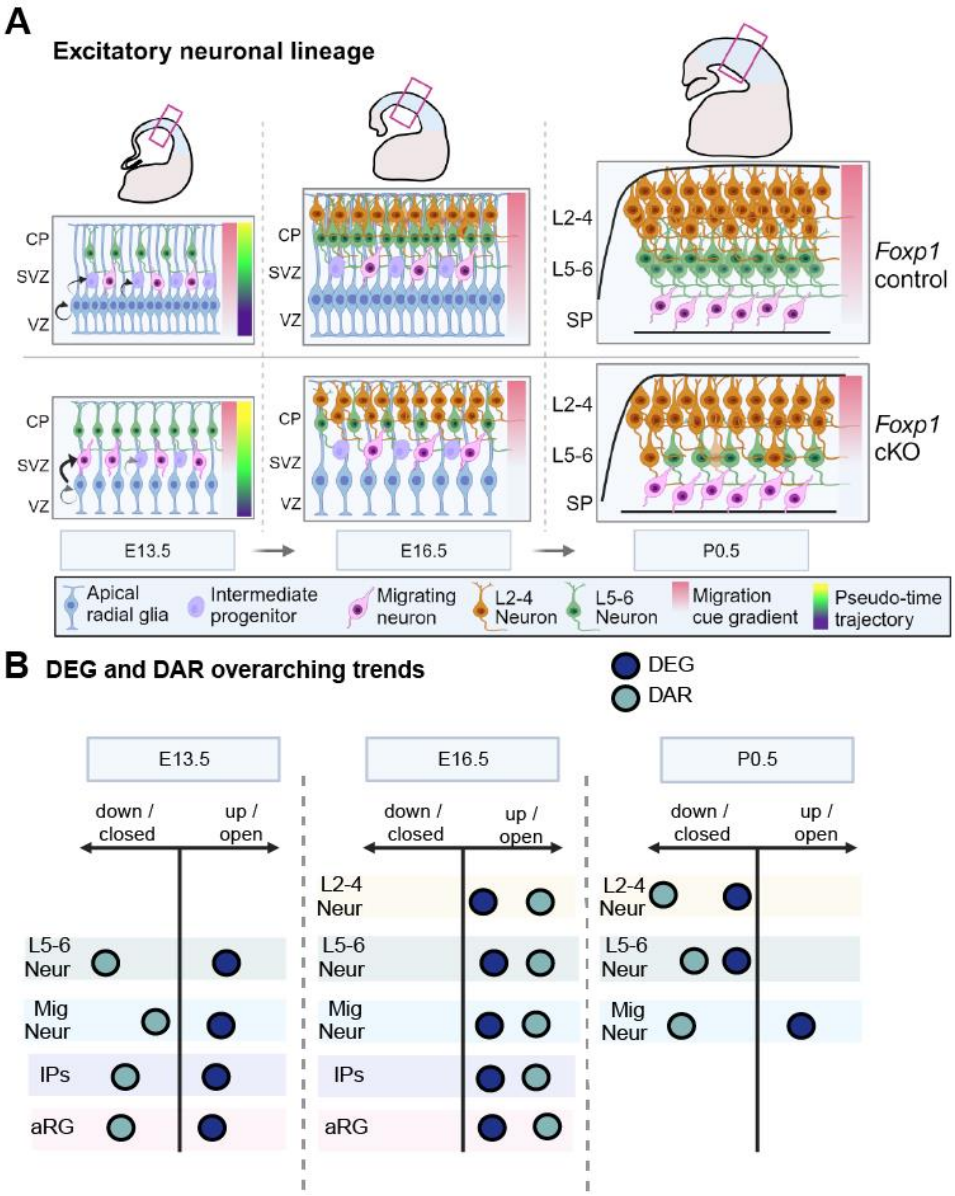

Figure 7: Schematic summary highlighting changes due to loss of *Foxp1* across development. A) E13.5: Increased neurogenic divisions bias precocious neuronal production, older pseudo-age, longer S-Phase and cell cycle length. E16.5: Faster cell-cycle exit biases UL neuronal production, but overall neurogenic output is reduced due to improper maintenance of progenitor population. Loss of *Foxp1* alters genes important for migration. P0.5: Total neuronal output is reduced in cKOs. Select UL neurons from progenitors dividing at E15.5 exhibit migrating deficit. *Foxp1* cKOs have relatively decreased DL neuron proportions and increased UL proportions compared to controls. **B)** Schematic represents overall changes in DEGs or DARs; at E13.5 there are more upregulated DEGs, but more closed DARs in cKOs. At E16.5, there is concordance in more upregulated DEGs and open DARs. At P0.5, there are more downregulated DEGs (expect in migrating neurons) but the difference between up and downregulated DEGs is modest. There are more closed DARs in all cell types.