

# Coordinated Epigenomic Control of Cortical Interneuron Lineage Specification and Ganglionic Eminence Diversity in Health and Disease

## INTRODUCTION:

Gamma-aminobutyric acid releasing or GABAergic cortical interneurons regulate local neural circuits and are critical for maintaining a precise excitatory-inhibitory balance for normal cortical function; imbalances due to GABAergic deficits have been linked to numerous psychiatric and neurodevelopmental disorders including epilepsy, schizophrenia, and autism. While much is still unknown about these highly complex disorders, genome-wide association studies (GWAS) have revealed that the majority of disease-associated SNPs are enriched in non-coding regions. In humans, the particular epigenetic mechanisms that drive cortical interneuron development are poorly understood. Recent collaborative efforts from the Shen and Kriegstein labs indicate that Sox2 maintains expression in cortical migrated interneurons and is implicated in regulation of development and subtype fate specification. In humans, the subcortical ganglionic eminence (GE) has been identified as the principal source for cortical interneurons. Characterization of Third Trimester Medial Ganglionic Eminence (MGE) has revealed an extended neurogenesis period of cortical interneurons. However, knowledge of cellular diversity remains limited. In collaboration with Eric Huang and Arnold Kriegstein, I have access to deidentified prefrontal cortex (PFC) and GE tissues from 2<sup>nd</sup> and 3<sup>rd</sup> Trimester. I will characterize Sox2 lineage specific activity during development across cortical interneurons subtypes from nuclei isolated (FANS) populations. I will explore the epigenomic diversity of the GE and identify disease-associated cis-regulatory elements. Through CRISPR mediated inhibition, the functional activity of such elements will be tested in primary culture. **I hypothesize that Sox2 and lineage-specific factors coordinate to modulate proper fate specification and development, and the disruption of active cis-regulatory elements due to disease-associated genetic variation could contribute to interneuron related disorders.**

## SPECIFIC AIMS:

### Aim I: Characterize Sox2 lineage specific activity during cortical interneuron development.

Sox2 is implicated in regulating lineage specific fate specification and remains expressed in cortical-migrated interneurons of the prefrontal cortex. However, systematic characterization of Sox2 regulation across GE populations has not been assessed. I will develop fluorescent-activated nuclei sorting (FANS) strategies to **A.** sort distinct cortical interneuron and progenitor populations across GE lineages. **B.** I will assess lineage-specific Sox2 regulation at gene promoters with CUT&TAG and RNA-seq. In collaboration with Dr. Tomasz Nowakowski, **C.** I will assess lineage-specific activities and dynamics of Sox2 cis-regulatory elements with integrated snATAC-seq. **D.** I will determine Sox2 co-occupancy with lineage-specific transcription factors using CUT&TAG.

### Aim II: Characterize disease associated cis-regulatory elements in diverse GE.

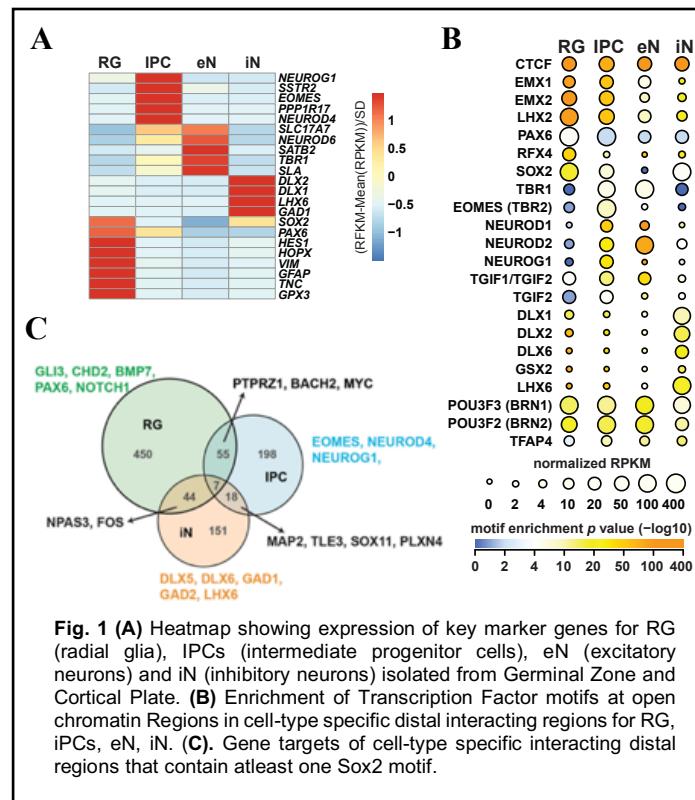
The GE produces almost all cortical interneurons in humans, and improper interneuron development is implicated in multiple neuropsychiatric diseases. However, the GE has been largely uncharacterized and much is unknown about its diverse cellular composition. While neuropsychiatric GWAS SNPs are known to be enriched in cis-regulatory elements in fetal tissues, chromatin accessibility of GE has not been assessed. **A.** I will map GE cellular diversity using single cell chromatin profiles **B.** I will assess GWAS SNPs disease association with cis-regulatory elements using HaploReg. **C.** In collaboration with Dr. Xiaoyu Yang of our lab, **D.** we will validate disease-associated candidate CRE in primary culture with CRISPRview.

**BACKGROUND:** Gamma-aminobutyric acid releasing or GABAergic cortical interneurons regulate local neural circuits and are critical for maintaining a precise excitatory-inhibitory balance for normal cortical function<sup>1,10</sup>, imbalances due to GABAergic deficits have been linked to numerous psychiatric and neurodevelopmental disorders including epilepsy, schizophrenia, and autism<sup>2,3</sup>. While human cortical expansion and the excitatory lineage have been well-characterized in the literature<sup>4,5</sup>, the molecular mechanisms that drive GABAergic interneuron development are poorly understood. The subcortical ganglionic eminence (GE) has been identified as the principal sources of cortical interneurons<sup>6,7</sup>. The medial ganglionic eminence (MGE) and caudal ganglionic eminence (CGE) generate distinct cortical interneuron subtypes through a successive developmental progression<sup>10</sup>. First, radial glia of the ventricular zone asymmetrically divide and produce intermediate progenitors (iPCs). In the subventricular zone, iPCs proliferate, and differentiate into newborn interneurons<sup>4,9</sup>. These interneurons mature and lose progenitor marks, migrating tangentially and then radially positioning into the cortex<sup>10</sup>.

Recent collaborative work from the Shen and Kriegstein labs<sup>21</sup> indicates that Sox2 maintains expression in MGE cortical migrated interneurons, yet downregulated in excitatory neurons (**Fig. 1A**). TF motif analysis (**Fig. 1B**) of cell-type specific non-coding regions has revealed enrichment for Sox2 binding, and in-silico gene targets include genes both necessary for proper interneuron development like DLX5, DLX6, GAD1, GAD2, and required for MGE cortical interneuron specification like Lhx6 (**Fig. 1C**)<sup>33</sup>. Additionally, Sox2 is required for MGE development and directly activates Nkx2.1<sup>32</sup>, and in cultured human MGE progenitors, Sox2 and Nkx2.1 co-associate at putative CREs<sup>68</sup>. Sox2 staining of 2<sup>nd</sup> and 3<sup>rd</sup> Trimester GE and single cell datasets of the developing cortex has revealed the distributed expression of Sox2 across neural progenitors, interneuron progenitors, and newborn interneurons. Sox2 acts as a master regulator in maintaining pluripotency in neural stem cells<sup>30,31</sup>; however, studies have also implicated its role in neurogenesis of neuron subtypes including hippocampal through priming the epigenetic landscape of pro-neural and neurogenic genes<sup>30,34</sup>. Altogether, these findings implicate a novel and extended role for Sox2 regulating cell maintenance, GABAergic interneuron maturation, and MGE subtype fate specification.

In terms of understanding transcription factor activity *in-vivo*, previous studies lack cell-type specific resolution and do not systematically follow cortical interneuron developmental progression to post-migration. While single-cell methods to profile protein-DNA binding exist like CUT&TAG<sup>29</sup>, it requires large quantities of limited custom Tn5 enzyme. We have successfully demonstrated a Fluorescent Activated-Cell Sorting (FACS) strategy to isolate cell populations from human developing cortex and conduct bulk epigenomic profiling<sup>21</sup>. Based on scRNA-seq data of the neocortex<sup>11</sup> and cell staining<sup>8,61-63</sup>, we have devised a novel approach to sort nuclei from Third Trimester PFC and GE tissue. With integrated single-nuclei ATAC-seq (snATAC-seq) analyses, we can recapitulate GE cortical interneuron development and achieve cell-type specific resolution of Sox2 binding *in-vivo*.

While Sox2 may play a key role in cortical interneuron development, studies in mouse have observed that cortical interneuron specification and migration result from precise spatiotemporal regional activity of multiple transcription factors in the MGE<sup>10,24-26</sup>. However, there is accumulating evidence of mouse and human species-specific differences of interneuron-related features<sup>6,17,27-28</sup>. Therefore, identifying



**Fig. 1** **(A)** Heatmap showing expression of key marker genes for RG (radial glia), IPCs (intermediate progenitor cells), eN (excitatory neurons) and iN (inhibitory neurons) isolated from Germinal Zone and Cortical Plate. **(B)** Enrichment of Transcription Factor motifs at open chromatin Regions in cell-type specific distal interacting regions for RG, iPCs, eN, iN. **(C)** Gene targets of cell-type specific interacting distal regions that contain atleast one Sox2 motif.

novel cis-regulatory mechanisms and transcription factors in human MGE will be essential to increment our understanding.

Recent characterization of Third Trimester MGE has revealed an extended neurogenesis of cortical interneurons until the end of pregnancy<sup>8</sup>. Previous work in this region has been sparse, identifying MGE progenitors populations based on transcription factor marker staining (**Fig. 2**)<sup>7,8</sup>.

However, both 1) cellular diversity of the GE and 2) molecular characterization of residing MGE cortical interneuron progenitors remain unknown.

Uncovering cellular diversity is critical for identifying possible glial and neuronal subpopulations affected due to preterm disorders that result in damaged GE, such as Germinal Matrix Hemorrhage<sup>58</sup>.

Previous studies have utilized single-cell RNA sequencing (scRNA-seq) to unravel the

transcriptomic diversity in developing brain tissues and in disease contexts<sup>11-14</sup>; however, these methods lack the propensity to identify gene regulatory mechanisms that drive expression differences. Alternatively, epigenetic features like chromatin accessibility are highly cell-type specific<sup>35-37</sup>, and changes in chromatin state are known to correspond with patterns of neuronal differentiation<sup>16-17,21,36,37</sup>. snATAC-seq has been able to identify lineage subtypes in developing and adult brain tissues, while simultaneously defining cell-type specific cis-regulatory elements and potentially active transcription factors<sup>16-18</sup>. By mapping cell-type specific open chromatin regions in GE, we can simultaneously explore cellular diversity and gain valuable insight into regulatory mechanisms of MGE cortical interneuron lineage specification.

Assessing the function of regulatory sequences is necessary to understand target effect on gene expression; however, validating identified cis-regulatory elements in their endogenous context has remained challenging, with most experiments utilizing cell lines or iPSC-derived cells<sup>19-20</sup>. Our lab has developed CRISPRview<sup>21</sup>, which allows for cell-type specific validation of cis-regulatory elements in heterogeneous primary cell culture by combining CRISPR<sup>22</sup>, RNAscope<sup>23</sup>, immunostaining, and image analysis<sup>57</sup>. By prioritizing cis-regulatory elements with co-localized epilepsy<sup>38</sup>, schizophrenia<sup>39</sup>, or autism<sup>40</sup> GWAS SNPs, we can validate disease-associated regulatory elements in their endogenous context and determine mechanisms of interneuron dysregulation.

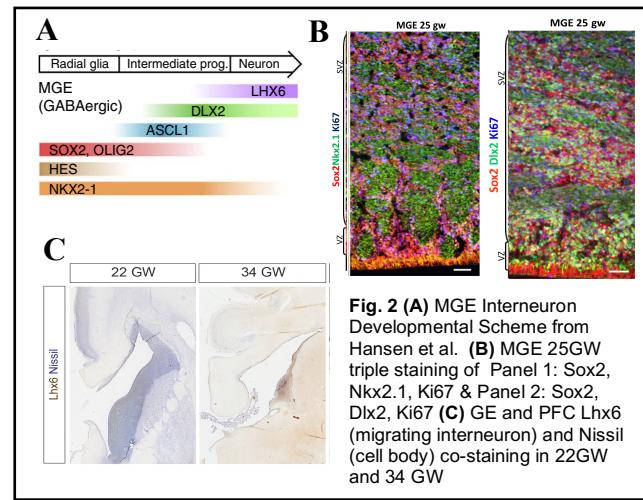
### Experimental Approach:

#### Aim I: Characterize shared and distinct cell-type specific Sox2 regulation during cortical interneuron development.

**Rationale:** Recent work has shown that Sox2 exhibits regulation of neurogenesis in a cell-type specific manner and could play an extended role in regulating both shared features of GABAergic interneuron development as well as MGE fate subtype specification (**Background & Fig. 1**) To modulate its diverse processes, Sox2 is known to partner with distinct transcription factors to confer its activities<sup>69,70</sup>. However, crucial cortical interneuron progenitors that reside in the GE have not been studied and Sox2 activity has not been directly assessed. By leveraging nuclei markers to sort GE and PFC populations and integrating snATAC-seq data to achieve cell-type specific resolution, we can understand the role of Sox2 and identify putative co-factors implicated in regulating GE cortical interneuron development. **I hypothesize that Sox2 regulates both GABAergic interneuron developmental and subtype fate specification in cortical interneurons through partnering with shared and distinct co-factors. (Schematic Fig. 3)**

#### Subaim IA. Sort nine cortical interneuron and progenitor populations from GE and PFC.

**FANS Nuclei Sorting:** I will adapt a previously established approach for fluorescence-activated cell sorting (FACS) strategy in our lab<sup>21</sup> and apply it to isolate nuclei from flash frozen tissue. The intact GE tissues range in age from 14GW to 30GW and are dorsal coronal slices that have LGE and MGE, but not CGE regions. I will purify nuclei from 50 mg each of cryostat sliced GE and PFC tissues using Minute Single Nucleus Isolation Kit for Neuronal Tissues / Cells (CAT #BN-020). To capture GE progenitors and

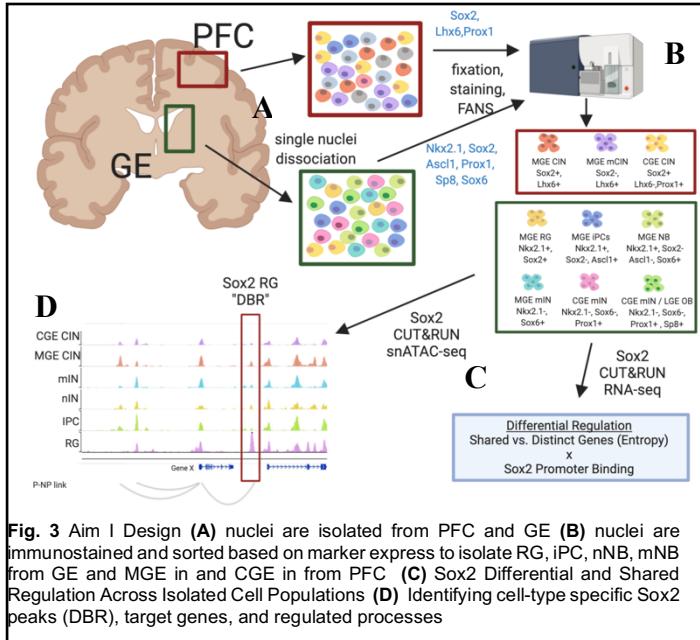


**Fig. 2** (A) MGE Interneuron Developmental Scheme from Hansen et al. (B) MGE 25GW triple staining of Panel 1: Sox2, Nkx2.1, Ki67 & Panel 2: Sox2, Dlx2, Ki67 (C) GE and PFC Lhx6 (migrating interneuron) and Nissl (cell body) co-staining in 22GW and 34 GW

interneurons, I will leverage known TF marker genes. Prox1 and Sp8 mark CGE cortical migrating interneurons that travel through the lateral migratory stream.<sup>63</sup> Sp8 also marks LGE derived olfactory bulb interneurons and progenitors.<sup>62</sup> Nkx2.1 is a canonical marker for MGE progenitors and Sox6 marks MGE-derived cortical migrating interneurons.<sup>62</sup>

To isolate GE migrating neurons and progenitors, I will fix nuclei and immunostain with NKX2.1-PA5-25940, Sox2-PerCP-Cy5.5, SP8-CL2707F, ASC11-sc-374104, SOX6-MBS460382, NBP1-30045 and sort nuclei as diagrammed on **Fig. 3**. Cortical staining of 3<sup>rd</sup> Trimester PFC shows that Sox2 is mainly restricted to interneurons and astrocytes. Based on scRNA-seq expression data<sup>11</sup>, Lhx6 and Prox1 mark MGE CIN and CGE CIN respectively. After cell fixation, nuclei isolation, and resuspension, I will incubate nuclei with Sox2-PerCP-Cy5.5, NBP1-30045, and LHX6-sc-271433 antibodies. I will sort nuclei as diagrammed on **Fig. 3** and enrich for MGE CIN and CGE CIN cell types.

**Expected Results:** I expect that using transcription factors will be suitable for isolating nuclei. I anticipate I will capture MGE radial glia (MGE RG), intermediate progenitors (MGE iPCs), newborn neuroblasts (MGE NB), migrating cortical interneurons (MGE mIN), Prox1+ CGE migrating interneurons (CGE mIN), and a mix SP8+ CGE interneurons and LGE olfactory bulb progenitors and neurons (CGE mIN / LGE OB). From the cortex, I anticipate I will capture MGE migrated cortical interneurons (MGE CIN), CGE migrated cortical interneurons (CGE CIN), and matured MGE CIN (MGE mCIN).



**Fig. 3 Aim I Design** (A) nuclei are isolated from PFC and GE (B) nuclei are immunostained and sorted based on marker express to isolate RG, iPC, nNB, mNB from GE and MGE in and CGE in from PFC (C) Sox2 Differential and Shared Regulation Across Isolated Cell Populations (D) Identifying cell-type specific Sox2 peaks (DBR), target genes, and regulated processes

#### Subaim IB: Characterize Sox2 shared and distinct promoter regulation across populations.

**Sox2 CUT&RUN:** I plan to conduct Sox2 CUT&RUN<sup>49</sup> in 150,000-200,000 cells of each PFC and GE isolated population using CUT&RUN Assay Kit #86652 with Sox2-PerCP-Cy5.5 antibody in triplicates for each sample. After library prep, I will sequence ~8-10 million 150bp paired-end reads per sample using Illumina NovaSeq, preprocess fastqs with fastp and align reads to hg38 genome using Bowtie2<sup>43</sup> with settings described in original CUT&RUN paper. I will call peaks using MACS2<sup>50</sup>, filter blacklisted regions<sup>52</sup>, and annotate peaks as described in Aim II. I will assess concordance of replicates and identify differentially bound regions across isolated populations using DiffBind<sup>45</sup>.

**RNA-seq:** I will conduct RNA-seq as previously described<sup>21</sup> using 100,000-150,000 cells to assess isolated population identity and gene expression. I will use DESeq2<sup>60</sup> to identify differentially expressed genes across all populations and identify population specific vs. shared genes by leveraging a method previously used<sup>21</sup> by measuring Shannon entropy across normalized gene expression counts. I will identify Sox2 regulated genes by intersecting Sox2 promoter bound sites with active genes.

**Expected Results:** Based on the role of Sox2 in stem cells, I expect to identify tens of thousands of Sox2 binding sites in MGE RG and iPCs populations, but I am open to the possibility of decreased binding in more mature MGE and CGE interneurons both pre- and post-migration. I expect Sox2 binding sites to lie in both promoter and non-coding regions. I anticipate that differential binding at promoters of Sox2 will correspond with differential gene expression patterns. Comparing across lineages, I expect Sox2 differential promoter binding to correspond with fate specifying genes and shared promoter binding to correspond to common interneuron developmental programs. I expect MGE progenitors to have Sox2 differential promoter binding at genomic loci associated with stem cell maintenance.

#### Subaim IC: Determine shared and distinct cell-type specific Sox2 co-factors and regulated biological processes by leveraging chromatin data.

**Integrating snATAC-seq data:** To match open chromatin data with isolated GE and PFC cell populations, I plan to integrate generated 3<sup>rd</sup> Trimester GE in Aim II with PFC snATAC-seq data generated in

collaboration with Tomasz Nowakowski. After batch correction and cluster identification using a similar pipeline as described in Aim II, I will determine open chromatin peaks per cluster by using MACS2. For each ATAC-seq cluster, I will identify Differentially Accessible Regions (DARs) and determine promoter to non-promoter (P-NP) links.

*Identifying Sox2 shared and differentially bound regions, HOMER TF analysis, GO analysis, :* To determine high-confidence cell-type specific Sox2 differentially bound regions (DBRs), I will first separate clusters based on marker genes for isolated cell populations. For each set of clusters, I will identify DBRs by looking at co-localization of chromatin and Sox2 peaks. For each cluster, I will conduct HOMER<sup>53</sup> TF analysis of Sox2 DBRs. Additionally, I will connect Sox2 DBR peaks to promoters of genes, and determine putative Sox2 regulated biological processes using Gene Ontology (GO)<sup>54,55</sup> term analysis. I will also compare shared Sox2 bound regions both across developmental “pseudotime” and by lineage.

**Expected Results:** I anticipate that Sox2 binding sites will mainly lie within accessible chromatin regions and therefore, partitioning Sox2 bound sites to cell clusters based on open chromatin will allow for cell-type specific resolution. I anticipate the majority of DBR Sox2 binding sites will coincide with DARs and lie in non-coding regions. I expect HOMER motif analysis of cell-type-specific Sox2 binding sites will reveal novel and known lineage-specific co-factors including Nkx2.1 for MGE progenitors. Comparing shared binding sites across developmental “pseudotime”, I expect to identify co-factors involved in interneuron development invariant of GE origin such as Ascl1, Dlx1, Dlx2 and others. I expect GO analysis to reveal gene patterns involved in both neuronal differentiation and cell maintenance for respective cell-types.

**Potential Caveats, Alternative and Future Directions:** FANS sorting relies on Flow Cytometry compatible antibodies. I have chosen markers based on the availability of commercially antibodies that have been validated for use with flow cytometry. However, if an antibody were to fail I would try multiple antibodies or adjust my sorting scheme to isolate fewer populations. As an alternative method to CUT&RUN, I can conduct single-cell CUT&TAG to achieve cell-type specific resolution of TF binding. While this aim will explore Sox2 driven patterns of gene regulation and putative co-regulators, follow-up to assess function will be necessary. Future studies should include co-immunoprecipitation or IP mass spectrometry assays to confirm interaction with identified co-factors and if an antibody is available, Re-ChIP<sup>67</sup> methods to confirm co-occupancy. This would most likely be infeasible in primary tissue; however, we could adapt in-vitro models of MGE cortical interneuron development<sup>64</sup>, or possibly organoid models that can recapitulate interneuron migration<sup>65</sup> for future validation. With in-vitro models, the transcriptome and epigenome would need to be assessed to understand biological similarity to primary tissues. Functional activity of Sox2 would also be key to confirm gene regulatory functions and could be flexibly introduced by replacing the Sox2 endogenous loci with an auxin inducible degron cassette<sup>66</sup>.

## **Aim II: Characterize cellular diversity and cis-regulatory element regulation of MGE cortical interneuron development in GE.**

**Rationale:** The ganglionic eminence has been identified as the principal source of cortical interneurons, and recent characterization of Third Trimester MGE has revealed an extended neurogenesis until the end of pregnancy<sup>8</sup>. However, characterization of GE cellular composition and mechanisms of MGE cortical interneuron development have remained limited. Creating a GE cellular atlas during this developmental period is critical for understanding the potential subpopulations affected by preterm disorders like Germinal Matrix Hemorrhage<sup>58</sup>. While single-cell transcriptomic methods have been leveraged to determine cellular identity in heterogeneous tissue<sup>11-13</sup>, single-cell chromatin profiling can simultaneously identify cell types and define cis-regulatory mechanisms that pertain to interneuron development<sup>16-18</sup>. Moreover, in developing brain tissues epigenomic features are incredibly informative, as changes in chromatin state drive neuronal differentiation<sup>21</sup>, preceding transcriptomic changes<sup>17</sup>. I hypothesize that cell-type specific cis-regulatory elements and transcription factors regulate genes that contribute to GE interneuron diversity and MGE cortical interneuron development. (Schematic Fig. 4)

### **Subaim IIA. Assess GE cellular diversity using single cell chromatin profiles.**

*Nuclei Isolation and 10X snATAC-seq library preparation:* I plan to conduct 10X snATAC-seq on GE tissue from 3 GW28-30 individuals and produce ~20,000 high quality nuclei profiles per sample. These intact GE tissues contain both lateral and medial ganglionic eminence, and for each sample, I will cryostat section multiple coronal slices. To assess quality and area of dissection, I will DAPI stain coronal-slices

and mark MGE by staining for Nkx2.1. I will isolate nuclei from 20 mg of tissue using Minute Single Nucleus Isolation Kit for Neuronal Tissues / Cells (CAT #BN-020). I will resuspend nuclei in PBS containing 5% BSA, count nuclei using Countess II Automated Cell Counter. 120,000-150,000 nuclei will be isolated, pelleted, and resuspended at 5,000 nuclei / uL for target 10,000 nuclei per lane. Nuclei will be loaded into two lanes and downstream 10X capture/library preparation will be conducted by IHG core following 10X ATAC GEM protocol.

**Cell Clustering:** I will use the SnapATAC<sup>41</sup> pipeline to cluster cells. SnapATAC improves on previous scATAC-seq analysis pipelines by measuring cell similarity without relying on a bulk reference accessibility profile which biases towards more represented cell types. I will use Latent Semantic Index to reduce dimensions and use scAlign<sup>42</sup> on latent dimensions to correct for batch effects across 10X ATAC runs. based on batch effect corrected latent dimensions, perform Leiden clustering, and visualize clusters using UMAP<sup>44</sup>.

**Annotating Clusters and peak calling:** I will calculate “gene activity scores” by creating a cell-by-gene matrix where each gene includes loci for gene-body, promoter, and 1kb upstream of the promoter. I will visually annotate clusters based on known marker genes. For each GE cell cluster, I will pool signal from all cells and call peaks using MACS2<sup>50</sup>.

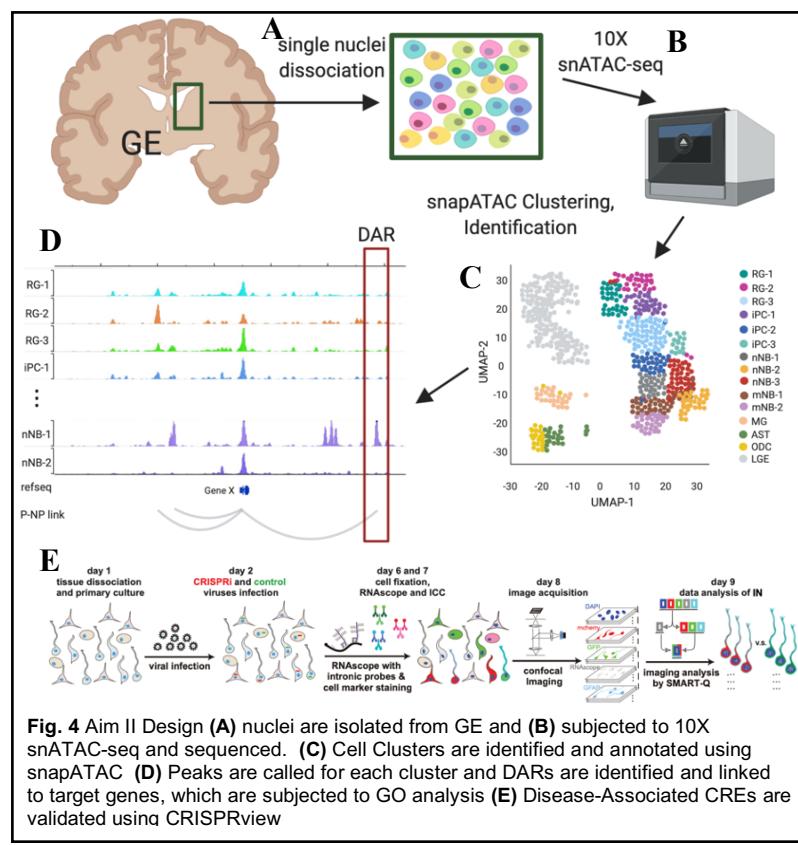
**Expected Results:** Our lab and others have shown that neuron and glial populations have generally distinct chromatin landscapes. I expect neuronal and glial populations to segregate when visually clustered. I anticipate based on Nkx2.1 “gene activity scores” I will be able to isolate MGE populations, and that cell clusters will segregate by region. I anticipate I will identify glial subtypes including astrocytes, oligodendrocytes, and microglia. In terms of neuronal populations, I expect to identify multiple radial glia, intermediate progenitor, and neuroblast populations including corresponding cortical interneuron clusters that will segregate based on developmental stage and interneuron subtype, yet resemble a lineage trajectory. Since progenitor populations are observed to have an accessible chromatin state, I expect to identify 70-100k peaks in each cell cluster.

#### Subaim IIB. Reconstruct cortical interneuron developmental trajectory and prioritize candidate transcription factor regulators with Cicero and HOMER.

**Constructing Developmental Trajectories:** I will use Cicero<sup>47</sup> to order lineage developmental trajectory using clusters with radial glia markers as “root cells”. Cicero applies reversed graph embedding to construct single-cell trajectories in an unsupervised manner.

**Annotating Differentially Accessible Elements and TF motif analysis:** To identify cell-type specific cis-regulatory elements, I will identify differentially accessible regions by performing a Two-Sided Fisher’s exact test with FDR < 0.05. To identify putative transcription factors (TF) that are involved in interneuron development, I will conduct TF binding motif enrichment analysis using HOMER for annotated non-coding DARs, and prioritize candidates by cross-assessing with TF “gene activity scores”.

**Expected Results:** I anticipate I will identify multiple lineage trajectories including both cortical interneuron and oligodendrocyte. I expect the majority of open chromatin regions and DARs to lie in non-coding regions. In cortical interneuron lineages, I anticipate that transcription factors that have enriched



**Fig. 4 Aim II Design** (A) nuclei are isolated from GE and (B) subjected to 10X snATAC-seq and sequenced. (C) Cell Clusters are identified and annotated using snapATAC (D) Peaks are called for each cluster and DARs are identified and linked to target genes, which are subjected to GO analysis (E) Disease-Associated CREs are validated using CRISPRview

To cluster cells, I will construct K-Nearest Neighbors graphs, perform Leiden clustering, and visualize clusters using UMAP<sup>44</sup>.

binding sites in DARs and high “gene activity scores” are important for regulating lineage specific processes including Dlx1, Dlx2, Sox2, Lhx6, Ascl1, among others.

### **Subaim IIC. Uncover distinct lineage-specific biological processes linked to chromatin state changes**

**Promoter to Non-Promoter (P-NP) linkage:** Co-accessibility of cis-regulatory elements and promoters has been shown to be a reliable predictor of DNA-DNA physical interaction<sup>18,47</sup>. Using a similar approach as described for DHS-seq, I will predict P-NP links based on the principle that interacting accessible elements will covary in their accessibility together across cell-types. I will create an aggregated peak-list from all cells and separate peaks into two lists based on promoter annotation. For each promoter, I will identify all non-promoter peaks within 500kb flanking region. For every possible P-NP defined, I will generate cluster count vectors filled by peak counts. I will conduct Pearson correlation of every P-NP paired vector and determine significant pairs by using a cutoff of >0.7.

**Gene Ontology analysis:** To determine differentially regulated genes, I will identify genes linked to DARs based on identified P-NP links. To uncover distinctly regulated biological processes, I will conduct Gene Ontology<sup>54,55</sup> analysis on genes identified for each cluster.

**Expected Results:** I anticipate that co-accessibility of P-NP links will predict regulatory activity. I expect that genes linked to DARs are differentially regulated and important for neuronal development and cell-type specific processes. I anticipate GO analysis will return neuro and glial related terms in appropriate cell types.

### **Subaim IID. Functionally validate disease associated cis-regulatory elements in primary culture.**

**Validation Design:** When live GE tissue is available, I plan to validate disease associated CRE using a previously established method in our lab, CRISPRview<sup>21</sup>. CRISPRview allows for cell-type specific validation of heterogeneous primary cell culture by combining CRISPRi, RNAscope, and immunostaining. We will identify candidate cis-regulatory elements through co-localization with epilepsy<sup>38</sup>, autism<sup>39</sup>, and schizophrenia<sup>40</sup> GWAS SNPs. We will design guides to target the center of open chromatin loci. In fresh tissues, cell types will be determined based on marker genes and cross-referenced with “gene activity scores” from snATAC-seq data. CRISPRview will be conducted as diagrammed (**Fig. 4E**). Briefly, after tissue dissociation, primary cultures of GE tissue will be transfected with two dCas9-KRAB containing lentivirus expressing 1) experimental sgRNA and mCherry and 2) control sgRNA and GFP. After five days in culture, cells will be fixed and stained using antibodies for mCherry, GFP, marker gene, and RNAscope (ACD+) probes. High resolution images will be taken using confocal microscopy, and the number of punctate dots representing nascent transcripts will be compared between experimental (mCherry+) and control (GFP+) sgRNA-treated cells. Image analysis will be conducted using SMART-Q<sup>57</sup> pipeline.

**Expected results:** I anticipate that GWAS SNPs localized in open chromatin regions will have associated activity in their respective cell type. I expect that CRISPRi inhibition of cis-regulatory elements will alter gene expression of linked target genes observed as a decrease in punctate dots as compared to control.

**Caveats, Alternative approaches, Future Directions:** “Gene activity scores” do not directly correlate with gene expression. To obtain gene expression for each cell-cluster, I could conduct 10X snRNA-seq in matched GE tissue and perform integrated analysis with snATAC-seq to generate pseudo multi-omics cells clusters in-silico. Conducting this complementary analysis could benefit defining cell-clusters, prioritizing transcription factors, and allow testing of additional hypotheses like the relationship between DARs and differentially expressed genes. Relying on a strict distance cutoff when identifying potential P-NP pairs could lead to potentially missed long distance P-NP interactions. Since almost all enhancer promoter interactions occur in topological associated domains (TADs)<sup>56</sup>, I could alternatively conduct Hi-C in nuclei dissociated from Third Trimester GE and call TADs. I could then revise my P-NP analysis to choose P-NP pairs based on localization within the same TAD. I could integrate this dataset to ask multi-layered epigenetic questions like exploring the spatial relationship of DARs in TADs. If available, an alternative method for functional validation would be to identify tissue regions or ages with more enriched cell populations of interest, and setup primary cultures with these enriched tissue. Then, it would be appropriate to treat tissue as bulk and use qPCR as readout instead of FISH.

### **CONCLUSION:**

The underlying goals of this work is to substantially advance our understanding of cell-type specific epigenetic mechanisms that contribute to interneuron subtype development and diversity in the greater

context of psychiatric and neurodevelopmental disease. Our first aim will elucidate the role of Sox2 in cortical interneuron lineage specification through isolating populations that follow the developmental trajectory. Our second aim will comprehensively profile the 3<sup>rd</sup> Trimester GE, identifying novel disease-associated cis-regulatory elements that are involved in regulating neuronal development. Future work could include follow up on identified transcription factors regulating GE cortical interneuron development, profiling additional time points, or incorporating additional epigenetic modalities such as methylation and 3D interaction to understand multi-layered mechanisms of developmental regulation.

## REFERENCES:

1. Ferguson, Brielle R., and Wen-Jun Gao. "PV Interneurons: Critical Regulators of E/I Balance for Prefrontal Cortex-Dependent Behavior and Psychiatric Disorders." *Frontiers in Neural Circuits*, vol. 12, 16 May 2018, 10.3389/fncir.2018.00037.
2. Powell, Elizabeth M., et al. "Genetic Disruption of Cortical Interneuron Development Causes Region- and GABA Cell Type-Specific Deficits, Epilepsy, and Behavioral Dysfunction." *The Journal of Neuroscience*, vol. 23, no. 2, 15 Jan. 2003, pp. 622–631, 10.1523/jneurosci.23-02-00622.2003.
3. Levy, L.M., and A.J. Degnan. "GABA-Based Evaluation of Neurologic Conditions: MR Spectroscopy." *American Journal of Neuroradiology*, vol. 34, no. 2, 19 Jan. 2012, pp. 259–265, 10.3174/ajnr.a2902.
4. Lui, Jan H., et al. "Development and Evolution of the Human Neocortex." *Cell*, vol. 146, no. 2, July 2011, p. 332, 10.1016/j.cell.2011.07.005.
5. Martínez-Cerdeño, Verónica, et al. "The Role of Intermediate Progenitor Cells in the Evolutionary Expansion of the Cerebral Cortex." *Cerebral Cortex*, vol. 16, no. suppl\_1, 1 July 2006, pp. i152–i161, 10.1093/cercor/bhk017.
6. Hansen, David V, et al. "Non-Epithelial Stem Cells and Cortical Interneuron Production in the Human Ganglionic Eminences." *Nature Neuroscience*, vol. 16, no. 11, 6 Oct. 2013, pp. 1576–1587, 10.1038/nn.3541.
7. Ma, Tong, et al. "Subcortical Origins of Human and Monkey Neocortical Interneurons." *Nature Neuroscience*, vol. 16, no. 11, 6 Oct. 2013, pp. 1588–1597, 10.1038/nn.3536.
8. Arshad, Arslan, et al. "Extended Production of Cortical Interneurons into the Third Trimester of Human Gestation." *Cerebral Cortex*, vol. 26, no. 5, 16 Apr. 2015, pp. 2242–2256, 10.1093/cercor/bhv074.
9. Brown, K. N., et al. "Clonal Production and Organization of Inhibitory Interneurons in the Neocortex." *Science*, vol. 334, no. 6055, 27 Oct. 2011, pp. 480–486, 10.1126/science.1208884.
10. Hu, Jia Sheng, et al. "Cortical Interneuron Development: A Tale of Time and Space." *Development*, vol. 144, no. 21, 31 Oct. 2017, pp. 3867–3878, 10.1242/dev.132852.
11. Nowakowski, Tomasz J., et al. "Spatiotemporal Gene Expression Trajectories Reveal Developmental Hierarchies of the Human Cortex." *Science*, vol. 358, no. 6368, 7 Dec. 2017, pp. 1318–1323, 10.1126/science.aap8809.
12. Ofengeim, Dmitry, et al. "Single-Cell RNA Sequencing: Unraveling the Brain One Cell at a Time." *Trends in Molecular Medicine*, vol. 23, no. 6, June 2017, pp. 563–576, 10.1016/j.molmed.2017.04.006.
13. Zhong, Suijuan, et al. "A Single-Cell RNA-Seq Survey of the Developmental Landscape of the Human Prefrontal Cortex." *Nature*, vol. 555, no. 7697, 2018, pp. 524–528, 10.1038/nature25980.
14. Velmeshev, Dmitry, et al. "Single-Cell Genomics Identifies Cell Type-Specific Molecular Changes in Autism." *Science*, vol. 364, no. 6441, 16 May 2019, pp. 685–689, 10.1126/science.aav8130.
15. Mathys, H., Davila-Velderrain, J., Peng, Z. et al. Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* 571, E1 (2019). <https://doi.org/10.1038/s41586-019-1329-6>.
16. Preissl, Sebastian, et al. "Single-Nucleus Analysis of Accessible chromatin in Developing Mouse Forebrain Reveals Cell-Type-Specific Transcriptional regulation." *Nature Neuroscience*, vol. 21, no. 7, 1 Mar. 2018, pp. 1015–1015, 10.1038/s41593-018-0114-4.
17. Ziffra, Ryan S, et al. "Single Cell Epigenomic Atlas of the Developing Human Brain and Organoids." *BioRxiv*, 1 Jan. 2019, p. 2019.12.30.891549
18. Ryan, Corces M, et al. "Singlecell Epigenomic Identification of Inherited Risk Loci in Alzheimer's and Parkinson's Disease." *BioRxiv*, 1 Jan. 2020, p. 2020.01.06.896159
19. Fulco, Charles P., et al. "Systematic Mapping of Functional Enhancer–Promoter Connections with CRISPR Interference." *Science*, vol. 354, no. 6313, 29 Sept. 2016, pp. 769–773, 10.1126/science.aag2445.
20. Song, Michael, et al. "Mapping Cis-Regulatory Chromatin Contacts in Neural Cells Links Neuropsychiatric Disorder Risk Variants to Target Genes." *Nature Genetics*, vol. 51, no. 8, 31 July 2019, pp. 1252–1262, 10.1038/s41588-019-0472-1.

21. Song, Michael, et al. "3D Epigenomic Characterization Reveals Insights Into Gene Regulation and Lineage Specification During Corticogenesis." *BioRxiv*, 1 Jan. 2020, p. 2020.02.24.963652, [biorxiv.org/content/early/2020/02/25/2020.02.24.963652.abstract](https://doi.org/10.1101/2020.02.24.963652), 10.1101/2020.02.24.963652.
22. Thakore, Pratiksha I., et al. "Highly Specific Epigenome Editing by CRISPR-Cas9 Repressors for Silencing of Distal Regulatory Elements." *Nature Methods*, vol. 12, no. 12, 26 Oct. 2015, pp. 1143–1149, 10.1038/nmeth.3630.
23. Wang, Fay, et al. "A Novel in Situ RNA Analysis Platform for Formalin-Fixed, Paraffin-Embedded Tissues." *J Mol Diagn*, vol. 14(1), no. 22–29, Jan. 2012, <https://dx.doi.org/10.1016%2Fj.jmoldx.2011.08.002>.
24. Xu, Qing, et al. "Fate Mapping Nkx2.1-Lineage Cells in the Mouse Telencephalon." *The Journal of Comparative Neurology*, vol. 506, no. 1, 1 Jan. 2008, pp. 16–29, 10.1002/cne.21529.
25. Long, Jason E., et al. "Dlx1&2 and Mash1 Transcription Factors Control MGE and CGE Patterning and Differentiation through Parallel and Overlapping Pathways." *Cerebral Cortex*, vol. 19, no. suppl\_1, 22 Apr. 2009, pp. i96–i106, 10.1093/cercor/bhp045.
26. Flandin, P., et al. "The Progenitor Zone of the Ventral Medial Ganglionic Eminence Requires Nkx2-1 to Generate Most of the Globus Pallidus But Few Neocortical Interneurons." *Journal of Neuroscience*, vol. 30, no. 8, 24 Feb. 2010, pp. 2812–2823, 10.1523/jneurosci.4228-09.2010.
27. Miller, Daniel J., et al. "Shared and Derived Features of Cellular Diversity in the Human Cerebral Cortex." *Current Opinion in Neurobiology*, vol. 56, June 2019, pp. 117–124, 10.1016/j.conb.2018.12.005.
28. Hodge, R.D., Bakken, T.E., Miller, J.A. et al. Conserved cell types with divergent features in human versus mouse cortex. *Nature* 573, 61–68 (2019). <https://doi.org/10.1038/s41586-019-1506-7>.
29. Kaya-Okur, Hatice S., et al. "CUT&TAG for Efficient Epigenomic Profiling of Small Samples and Single Cells." *Nature Communications*, vol. 10, no. 1, 29 Apr. 2019, 10.1038/s41467-019-09982-5.
30. Kamachi, Y., and H. Kondoh. "Sox Proteins: Regulators of Cell Fate Specification and Differentiation." *Development*, vol. 140, no. 20, 1 Oct. 2013, pp. 4129–4144, 10.1242/dev.091793.
31. Bertolini, Jessica A., et al. "Mapping the Global Chromatin Connectivity Network for Sox2 Function in Neural Stem Cell Maintenance." *Cell Stem Cell*, vol. 24, no. 3, Mar. 2019, pp. 462–476.e6, 10.1016/j.stem.2019.02.004.
32. Ferri, A., et al. "Sox2 Is Required for Embryonic Development of the Ventral Telencephalon through the Activation of the Ventral Determinants Nkx2.1 and Shh." *Development*, vol. 140, no. 6, 26 Feb. 2013, pp. 1250–1261, 10.1242/dev.073411.
33. Liodis, P., et al. "Lhx6 Activity Is Required for the Normal Migration and Specification of Cortical Interneuron Subtypes." *Journal of Neuroscience*, vol. 27, no. 12, 21 Mar. 2007, pp. 3078–3089, 10.1523/jneurosci.3055-06.2007.
34. Amador-Arjona, Alejandro, et al. "SOX2 Primes the Epigenetic Landscape in Neural Precursors Enabling Proper Gene Activation during Hippocampal Neurogenesis." *Proceedings of the National Academy of Sciences*, vol. 112, no. 15, 30 Mar. 2015, pp. E1936–E1945, 10.1073/pnas.1421480112.
35. Thurman, R., Rynes, E., Humbert, R. et al. The accessible chromatin landscape of the human genome. *Nature* 489, 75–82 (2012). <https://doi.org/10.1038/nature11232>.
36. de la Torre-Ubieta, Luis, et al. "The Dynamic Landscape of Open Chromatin during Human Cortical Neurogenesis." *Cell*, vol. 172, no. 1–2, Jan. 2018, pp. 289–304.e18, 10.1016/j.cell.2017.12.014.
37. Mo, Alisa, et al. "Epigenomic Signatures of Neuronal Diversity in the Mammalian Brain." *Neuron*, vol. 86, no. 6, June 2015, pp. 1369–1384, 10.1016/j.neuron.2015.05.018.
38. The International League Against Epilepsy Consortium on Complex Epilepsies. "Genome-Wide Mega-Analysis Identifies 16 Loci and Highlights Diverse Biological Mechanisms in the Common Epilepsies." *Nature Communications*, vol. 9, no. 1, Dec. 2018, 10.1038/s41467-018-07524-z.
39. Huo, Yongxia, et al. "Functional Genomics Reveal Gene Regulatory Mechanisms Underlying Schizophrenia Risk." *Nature Communications*, vol. 10, no. 1, 8 Feb. 2019, 10.1038/s41467-019-08666-4.
40. Grove, J., Ripke, S., Als, T.D. et al. Identification of common genetic risk variants for autism spectrum disorder. *Nat Genet* 51, 431–444 (2019). <https://doi.org/10.1038/s41588-019-0344-8>

41. Fang, Rongxin, et al. "Fast and Accurate Clustering of Single Cell Epigenomes Reveals Cis-Regulatory Elements in Rare Cell Types." *BioRxiv*, 1 Jan. 2019, p. 615179, [biorxiv.org/content/early/2019/05/13/615179.abstract](https://www.biorxiv.org/content/early/2019/05/13/615179.abstract), 10.1101/615179.
42. Johansen, Nelson, and Gerald Quon. "ScAlign: A Tool for Alignment, Integration, and Rare Cell Identification from ScRNA-Seq Data." *Genome Biology*, vol. 20, no. 1, 14 Aug. 2019, 10.1186/s13059-019-1766-4.
43. Langmead, Ben, and Steven L Salzberg. "Fast Gapped-Read Alignment with Bowtie 2." *Nature Methods*, vol. 9, no. 4, 4 Mar. 2012, pp. 357–359, 10.1038/nmeth.1923.
44. McInnes, Leland, et al. "UMAP: Uniform Manifold Approximation and Projection." *Journal of Open Source Software*, vol. 3, no. 29, 2 Sept. 2018, p. 861, 10.21105/joss.00861.
45. Stark R, Brown G (2011). *DiffBind: differential binding analysis of ChIP-Seq peak data*. <http://bioconductor.org/packages/release/bioc/vignettes/DiffBind/inst/doc/DiffBind.pdf>
46. Nott, Alexi, et al. "Brain Cell Type-Specific Enhancer-Promoter Interactome Maps and Disease Risk Association." *Science*, vol. 366, no. 6469, 29 Nov. 2019, p. 1134, [science.sciencemag.org/content/366/6469/1134.abstract](https://science.sciencemag.org/content/366/6469/1134.abstract), 10.1126/science.aay0793.
47. Pliner, Hannah A., et al. "Cicero Predicts Cis-Regulatory DNA Interactions from Single-Cell Chromatin Accessibility Data." *Molecular Cell*, vol. 71, no. 5, Sept. 2018, pp. 858-871.e8, 10.1016/j.molcel.2018.06.044.
48. Sandberg, Magnus, et al. "Transcriptional Networks Controlled by NKX2-1 in the Development of Forebrain GABAergic Neurons." *Neuron*, vol. 91, no. 6, Sept. 2016, pp. 1260–1275, 10.1016/j.neuron.2016.08.020.
49. Skene, Peter J, and Steven Henikoff. "An Efficient Targeted Nuclease Strategy for High-Resolution Mapping of DNA Binding Sites." *ELife*, vol. 6, 16 Jan. 2017, 10.7554/elife.21856. Accessed 22 Mar. 2020.
50. Zhang, Yong, et al. "Model-Based Analysis of ChIP-Seq (MACS)." *Genome Biology*, vol. 9, no. 9, 2008, p. R137, 10.1186/gb-2008-9-9-r137.
51. [biorxiv.org/content/early/2019/12/31/2019.12.30.891549.abstract](https://www.biorxiv.org/content/early/2019/12/31/2019.12.30.891549.abstract), 10.1101/2019.12.30.891549.
52. Amemiya, Haley M., et al. "The ENCODE Blacklist: Identification of Problematic Regions of the Genome." *Scientific Reports*, vol. 9, no. 1, 27 June 2019, 10.1038/s41598-019-45839-z.
53. Heinz S, Benner C, Spann N, Bertolino E et al. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol Cell* 2010 May 28;38(4):576-589. PMID: [20513432](#)
54. Ashburner et al. Gene ontology: tool for the unification of biology. *Nat Genet*. May 2000;25(1):25-9.
55. The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res*. Jan 2019;47(D1):D330-D338.
56. Krijger, Peter Hugo Lodewijk, and Wouter de Laat. "Regulation of Disease-Associated Gene Expression in the 3D Genome." *Nature Reviews Molecular Cell Biology*, vol. 17, no. 12, 9 Nov. 2016, pp. 771–782, 10.1038/nrm.2016.138.
57. Bergenholz, Seth, et al. "SMARTQ: An Integrative Pipeline Quantifying Cell Type-Specific RNA Transcription." *BioRxiv*, 1 Jan. 2020, p. 2020.01.24.917872, [biorxiv.org/content/early/2020/01/24/2020.01.24.917872.abstract](https://www.biorxiv.org/content/early/2020/01/24/2020.01.24.917872.abstract), 10.1101/2020.01.24.917872.
58. Volpe, J. J. "Impaired Neurodevelopmental Outcome After Mild Germinal Matrix-Intraventricular Hemorrhage." *PEDIATRICS*, vol. 136, no. 6, 23 Nov. 2015, pp. 1185–1187, 10.1542/peds.2015-3553.
59. Yu, G., Wang, L. G. & He, Q. Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* 31, 2382-2383, doi:10.1093/bioinformatics/btv145 (2015).
60. Love, Michael I, et al. "Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2." *Genome Biology*, vol. 15, no. 12, Dec. 2014, 10.1186/s13059-014-0550-8.
61. Ma, T., et al. "A Subpopulation of Dorsal Lateral/Caudal Ganglionic Eminence-Derived Neocortical Interneurons Expresses the Transcription Factor Sp8." *Cerebral Cortex*, vol. 22, no. 9, 20 Oct. 2011, pp. 2120–2130, 10.1093/cercor/bhr296.

62. Ma, Tong, et al. "Subcortical Origins of Human and Monkey Neocortical Interneurons." *Nature Neuroscience*, vol. 16, no. 11, 6 Oct. 2013, pp. 1588–1597, 10.1038/nn.3536.
63. Rubin, Anna Noren, and Nicoletta Kessaris. "PROX1: A Lineage Tracer for Cortical Interneurons Originating in the Lateral/Caudal Ganglionic Eminence and Preoptic Area." *PLoS ONE*, vol. 8, no. 10, 14 Oct. 2013, p. e77339, 10.1371/journal.pone.0077339.
64. Close, Jennie L et al. "Single-Cell Profiling of an In Vitro Model of Human Interneuron Development Reveals Temporal Dynamics of Cell Type Production and Maturation." *Neuron* vol. 93,5 (2017): 1035-1048.e5. doi:10.1016/j.neuron.2017.02.014
65. Bagley, J., Reumann, D., Bian, S. et al. Fused cerebral organoids model interactions between brain regions. *Nat Methods* **14**, 743–751 (2017).  
<https://doi.org/10.1038/nmeth.4304>
66. Li, S., Prasanna, X., Salo, V.T. et al. An efficient auxin-inducible degron system with low basal degradation in human cells. *Nat Methods* **16**, 866–869 (2019).  
<https://doi.org/10.1038/s41592-019-0512-x>
67. Truax, Agnieszka D., and Susanna F. Greer. "ChIP and Re-ChIP Assays: Investigating Interactions Between Regulatory Proteins, Histone Modifications, and the DNA Sequences to Which They Bind." *Methods in Molecular Biology*, 22 Oct. 2011, pp. 175–188, 10.1007/978-1-61779-376-9\_12.
68. Goulburn, Adam L., et al. "Generating GABAergic Cerebral Cortical Interneurons from Mouse and Human Embryonic Stem Cells." *Stem Cell Research*, vol. 8, no. 3, May 2012, pp. 416–426, 10.1016/j.scr.2011.12.002.
69. Kamachi, Y., and H. Kondoh. "Sox Proteins: Regulators of Cell Fate Specification and Differentiation." *Development*, vol. 140, no. 20, 1 Oct. 2013, pp. 4129–4144, 10.1242/dev.091793.
70. Hagey, Daniel W., et al. "SOX2 Regulates Common and Specific Stem Cell Features in the CNS and Endoderm Derived Organs." *PLOS Genetics*, vol. 14, no. 2, 12 Feb. 2018, p. e1007224, 10.1371/journal.pgen.1007224.

Schematics created using BioRender.