Spatial enhancer activation determines inhibitory neuron identity

Elena Dvoretskova[1,2,+], May C. Ho[1,2,+], Volker Kittke[3,4,+], Ilaria Vitali[2], Daniel D. Lam[3,4], Irene Delgado[5], Chao Feng[1,2], Miguel Torres[5], Juliane Winkelmann[3,4,6] and Christian Mayer[1,2,✉]

[1]Max Planck Institute for Biological Intelligence, Martinsried, 82152, Germany
[2]Max Planck Institute of Neurobiology, Martinsried, 82152, Germany
[3]Helmholtz Zentrum München GmbH, German Research Center for Environmental Health, Institute of Neurogenomics, Neuherberg, 85764 Germany
[4]Technical University of Munich, TUM School of Medicine, Institute of Human Genetics, Munich, 81675, Germany
[5]Centro Nacional de Investigaciones Cardiovasculares (CNIC), Cardiovascular Development Program, Madrid, 28029, Spain
[6]Munich Cluster for Systems Neurology (SyNergy), Munich, 81377, Germany
[+]these authors contributed equally to this work
✉christian.mayer@bi.mpg.de

Monday 30th January, 2023

Abstract

The mammalian telencephalon contains a tremendous diversity of GABAergic projection neuron and interneuron types, that originate in a germinal zone of the embryonic basal ganglia. How genetic information in this transient structure is transformed into different cell types is not yet fully understood. Using a combination of *in vivo* lineage tracing, CRISPR perturbation and ChIP-seq in mice, we found that the transcription factor MEIS2 favors the development of projection neurons through genomic binding sites in regulatory enhancers of projection neuron specific genes. MEIS2 requires the presence of the homeodomain transcription factor DLX5 to direct its functional activity towards these sites. In interneurons, the activation of projection neuron specific enhancers by MEIS2 and DLX5 is repressed by the transcription factor LHX6. When MEIS2 carries a mutation associated with intellectual disability in humans, it is less effective at activating enhancers involved in projection neuron development. This suggests that GABAergic differentiation may be impaired in patients carrying this mutation. Our research has uncovered a mechanism by which the selective activation of enhancers plays a crucial role in the establishment of neuronal identity, as well as in potential pathological mechanisms.

Introduction

The ganglionic eminences (GEs) are embryonic subpallial structures that give rise to GABAergic projection neurons (PNs) of the striatum, globus pallidus, and amygdala, as well as to GABAergic interneurons (INs) of the cortex, hippocampus, and striatum (Bandler et al., 2017). It is divided into three spatial regions: the medial (MGE), caudal (CGE), and lateral (LGE) ganglionic eminences (Wonders and Anderson, 2006; Gelman et al., 2011; Anderson et al., 2001). Several transcription factors (TFs) and their co-factors have been shown to be necessary for the specification of GABAergic subtypes. (Leung et al., 2022; Flames et al., 2007), and their dysregulation results in disease (Leung et al., 2022; Zug, 2022). For example, members of the DLX family are present in the GE and are required for the development of GABAergic neurons (Anderson et al., 1997; Stühmer et al., 2002; Lindtner et al., 2019). The LIM homeodomain TF LHX6 is one of the factors known to regulate the generation of MGE-derived INs (Sandberg et al., 2016; Zhao et al., 2008), whereas MEIS2, a member of the TALE family of homeodomain-containing transcription factors, has been implicated in the generation of LGE-derived GABAergic PNs (Su et al., 2022). Haploinsufficiency of *MEIS2* in humans results in cardiac and palate abnormalities, developmental delay, and intellectual disability (Louw et al., 2015; Douglas et al., 2018; Giliberti et al., 2020; Zhang et al., 2021). The mechanisms by which these TFs select and activate their targets remain unclear.

Here, we used sparse CRISPR/Cas-mediated perturbation of *Meis2*, *Lhx6* and *Tcf4* in GABAergic progenitors and tracked their developmental trajectories with lineage barcodes and single-cell RNA sequencing (scRNA-seq). We found that the sparse perturbation of *Meis2* in the GE deranges the development of GABAergic neurons, increasing the proportion of IN clones at the expense of PN clones. We identified genomic target sites of MEIS1/2 in enhancers of PN-specific genes, through which MEIS2 induces PN fate. These sites frequently overlapped with binding sites of DLX5 and some also overlapped with target sites of LHX6. We performed luciferase reporter assays and found that DLX5 could activate these enhancers to some extent, but that MEIS2 was additionally required for strong activation. LHX6 repressed the DLX5/MEIS2-induced cooperative

activation of PN gene enhancers to promote an IN fate. Finally, a mutation of *Meis2* that causes intellectual disability in humans (Giliberti et al., 2020; Gangfuß et al., 2021) was much less able to potentiate the DLX5-induced activation of these enhancers. Our results indicate that MEIS2 acts as a transcriptional activator to generate patterns of enhancer activation that specifies PN identities within GABAergic precursor cells. This mechanism may contribute to neurological dysfunction in diseases caused by *MEIS2* mutations.

Results

In vivo tCROP-seq to assess the function of MEIS2 during GABAergic fate decisions

To investigate the effects of MEIS2 perturbation on cellular fate decisions in a sparse population of precursors in the GE, we modified CROP-seq (Datlinger et al., 2017), a method for pooled CRISPR screens with single-cell transcriptome readout. Instead of lentiviral vectors to deliver single-guide RNAs (sgRNAs), we used a PiggyBac **t**ransposon-based strategy (**t**CROP-seq) and *in utero* electroporation to efficiently deliver sgRNAs to cycling progenitors in the GE. The transposon system allows genes to be stably integrated into the genomes of electroporated cells and thus to be transmitted to their postmitotic daughter cells (Ding et al., 2005). This increases the pool of perturbed cells and ensures that the perturbation occurs during a period covering the peak of neurogenesis (Bandler et al., 2022). We also added specific capture sequences to the sgRNA vectors that efficiently link sgRNAs to cell barcodes, and enable sequencing of the protospacer from the transcriptome (Replogle et al., 2020). tCROP-seq sgRNA vectors also encode TdTomato to enable the labeling and enrichment of perturbed neurons. The efficiency of sgRNAs to induce frame-shift mutations was validated *in vitro* prior to the tCROP-seq experiments (Table S1).

The tCROP-seq vectors were targeted by *in utero* electroporation at E12.5 to progenitor cells of the GE in a mouse line ubiquitously expressing Cas9 (Platt et al., 2014) (Figure 1a). We

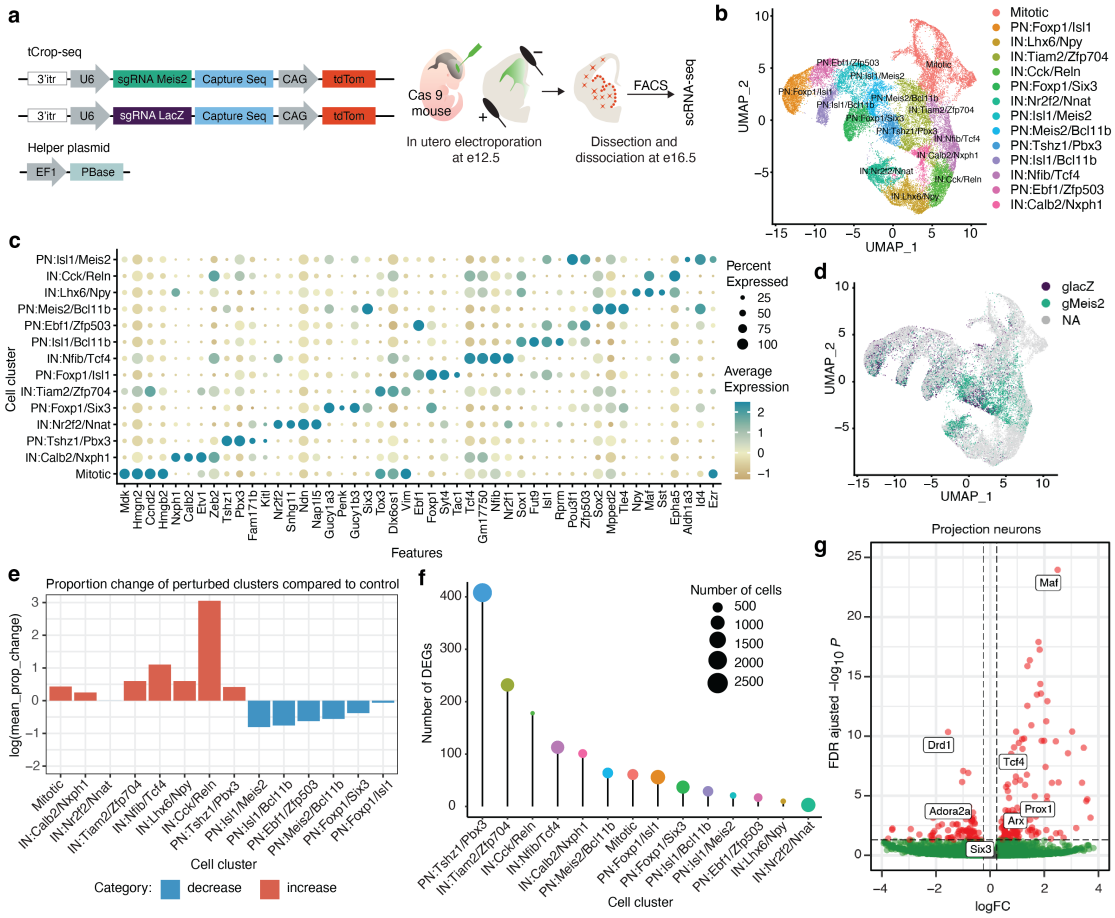


Figure 1: *in vivo* tCROP-seq of *Meis2* in the mouse forebrain. **a**, Vector maps and schematic of the *in vivo* tCROP-seq workflow, in which mutations in individual genes are introduced *in utero* and the effect is determined at a later time point via scRNA-seq. **b**, Uniform Manifold Approximation and Projection (UMAP) plot of inhibitory cells colored by clusters. **c**, Dotplot of the top five marker genes of inhibitory clusters. **d**, UMAP plot of the integrated dataset colored by sgRNA. **e**, Relative increase or decrease in the number of inhibitory cell clusters in gMeis2 compared to gLacZ. **f**, Lollipop plots showing the impact of gMeis2 on inhibitory clusters. **g**, Volcano plot depicting differentially expressed genes in gMeis2 and gLacZ projection neurons.

electroporated a total of 14 embryos from multiple pregnant females. Of these, 8 received sgRNAs for *Meis2* (gMeis2) and 6 received sgRNAs for *LacZ* (gLacZ), which served as a control. At E16.5, most TdTomato+ cells had migrated away from the ventricular zone and colonized a variety of structures, including the striatum, cerebral cortex, and olfactory bulb (Figure S1a), consistent with the migration patterns of GE-derived inhibitory neurons at this stage (Anderson et al., 2001). Cortices, striata, and olfactory bulbs were then dissected and TdTomato+ cells were enriched by FACS. tCROP-seq allows the retrospective assessment of which sgRNA was expressed in which cell. We pooled cells from embryos having received gLacZ or gMeis2, and conducted multiplexed single-cell RNA sequencing to minimize batch effects (Figure 1a; see Methods) (Jin et al., 2020). We sequenced 6 independent scRNA-seq experiments. Together, this resulted in a dataset containing 34481 cells passing quality controls and filtering, that could be linked with either gLacZ (11009) or gMeis2 (23472). We projected cells into a shared embedding using Harmony (Korsunsky et al., 2019) and applied a standard Seurat pipeline (Figure S2a).

Single perturbation of MEIS2 alters the proportion of PNs and INs

Louvain clustering grouped glia cells, excitatory neurons, and inhibitory neurons into multiple clusters (Figure S2a). We subset cells from inhibitory clusters (Figure S2b-e) and integrated them with scRNA-seq datasets from wild-type mice (Bandler et al., 2022), to get a higher resolution of inhibitory cell states (Figure1b). We annotated 14 inhibitory clusters based on shared marker gene expression and grouped them into three major classes: mitotic (mitotic), GABAergic PNs (PN:Foxp1/Six3, PN:Foxp1/Isl1, PN:Isl1/Bcl11b, PN:Ebf1/Zfp503, PN:Meis2/Bcl11b, PN:Isl1/Meis2, PN:Tshz1/Pbx3), and GABAergic INs (IN:Calb2/Nxph1, IN:Tiam2/Zfp704, IN:Nfib/Tcf4, IN:Lhx6/Npy, IN:Cck/Reln, IN:Nr2f2/Nnat; Figure 1b-c, Table S2). Cells expressing gMeis2 contained a reduced proportion of PN cell-types and an increased proportion of IN cell-types, when compared to gLacZ controls (Figure 1d-e, Figure S1b). This suggests that, under normal conditions, MEIS2 promotes the generation of PNs at the expense of INs. A pseudo-bulk differential gene expression analysis (DEG) (Squair et al., 2021) of GABAergic neurons comparing gMeis2 and

gLacZ showed reduced expression levels of genes known to be involved in PN development and increased expression levels of genes known to be involved in IN development (Table S3). The impact of gMeis2 on differential gene expression was strongest on immature clusters: PN:Tshz1/Pbx3 and IN:Tiam2/Zfp704 (Figure 1f, Table S4). In PN clusters, gMeis2+ cells showed decreased expression levels of genes known to be associated with PN identity, such as *Gucy1a3*, *Adora2a*, *Drd1*, *Six3*, and *Zfp503* (Kreitzer and Malenka, 2008; Song et al., 2021; Shang et al., 2022; Knowles et al., 2021), compared to gLacZ (Figure 1g, Table S3-4). Surprisingly, many genes related to IN development and specification, such as *Maf*, *Tcf4*, *Prox1*, *Arx*, *Sp8*, *Npas1* and *Nxph1* (Lim et al., 2018; Miyoshi et al., 2015; Batista-Brito et al., 2008; Wei et al., 2019) were up-regulated in PN clusters (Figure 1g, Table S3-4). In addition, GO Term analysis of the up and down-regulated DEGs reveal that processes such as neuron development, axon extension, and neuron differentiation are deregulated (Figure S2f). This raises the question of how neurons with a broad PN identity (Louvain clustering grouped them into PNs) acquired CGE/MGE-IN signatures. One possibility would be that, upon the perturbation of gMeis2, progenitors of the LGE-PN lineage fail to establish proper PN identity and switch to a CGE/MGE-IN identity.

Combined *in vivo* lineage tracing and tCROP-Seq reveal a shift in clonal compositions of perturbed cells

To test this possibility, we combined tCROP-seq with a barcode lineage tracing method called TrackerSeq (Bandler et al., 2022)), that integrates DNA barcodes into the genome of electroporated mitotic progenitors, enabling the tracking of clonal relationships between their postmitotic daughter neurons (Figure 2a). tCROP-seq and TrackerSeq can be used simultaneously because we have implemented a similar transposase strategy for both methods (Figure 2a). If a fate switch occurred in the presence of gMeis2, we would expect to observe a shift in clonal compositions from PNs to INs. We used *in utero* electroporation at E12.5 to introduce the TrackerSeq barcode library and tCROP-seq sgRNAs to cycling progenitors in the GE. We collected tdTomato/EGFP+ cells from 4 independent batches and prepared sequencing libraries for transcriptomes, sgRNAs, and

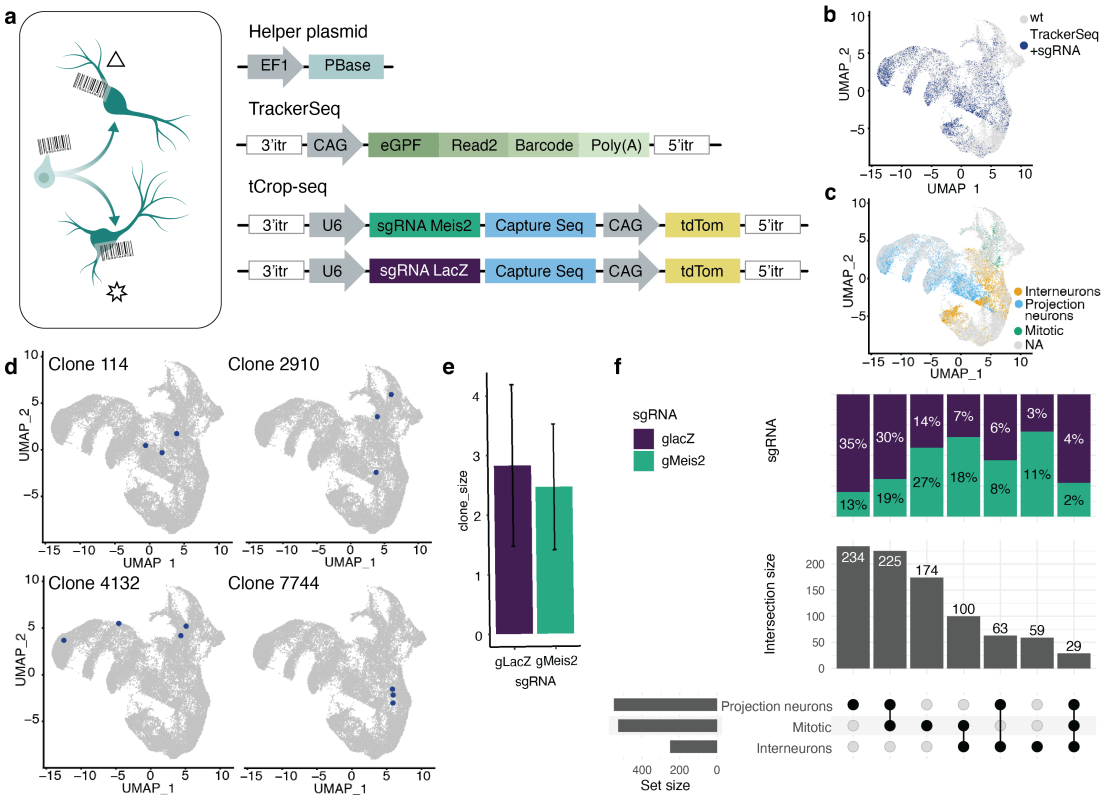


Figure 2: ***In vivo*** **TrackerSeq lineage tracing and tCROP-seq perturbation of** ***Meis2*****.** **a**, Schematic of the TrackerSeq lineage tracing process, in which clonal boundaries are determined using a diverse library of RNA tags. **b**, UMAP of the integrated dataset where cells that contained TrackerSeq lineage barcodes are colored. **c**, UMAP of the integrated dataset colored by cell class (mitotic, interneurons, projection neurons). **d**, Examples of clones that are shared between classes, and an example of a clone restricted to one class. **e**, Bar graph depicting the average clone size of inhibitory clones in the glacZ and gMeis2 datasets. **f**, UpSet plot showing clonal intersections between cell classes. The bar graph on top displays the proportion of clones belonging to gLacZ or gMeis2, the bar graph in the middle shows the number of observed intersections, and the bar graph on the left indicates the number of cells per cluster.

lineage barcodes. The cells with TrackerSeq barcodes were already part of the preceding tCROP-seq analysis and were thus integrated in the same embedding (Figure 2b). Consistent with Bandler et al. 2022 (Bandler et al., 2022), we found clones composed of mitotic cells, PNs, INs, and combinations thereof (Figure 2c-d). The clonal size of multi-cell clones was unchanged in gMeis2 compared to gLacZ (Figure 2e). The proportion of clones consisting of only mitotic cells was increased in gMeis2 compared to gLacZ, which agrees with a recent report showing that MEIS2 is required for LGE progenitors to leave the cell cycle (Su et al., 2022) (Figure 2f). We found many clones that dispersed across cell states. For example, 225 clones consisted of mitotic cells and PNs (mitotic-PN), and 100 clones consisted of mitotic cells and INs (mitotic-IN; Figure 2f). Strikingly, when we compared clonal patterns of gMeis2 and gLacZ cells, we observed a pronounced shift toward IN-only and mitotic-IN clones. Conversely, the number of PN-only, and mitotic-PN clones was decreased (Figure 2f). Our results suggest that perturbation of cells with gMeis2 causes a fate switch from PNs to INs.

Genomic binding of DLX5 and MEIS2 in the embryonic GE

To identify direct target genes of MEIS2, we performed chromatin immunoprecipitation-sequencing (ChIP-seq) on GE tissue dissected from E14.5 mouse embryos, using a combination of anti-MEIS1/2 and anti-MEIS2 antibodies. In the GE, the expression of *Meis2* is higher and more widespread than that of *Meis1*, therefore the antibodies are likely to bind primarily to MEIS2 epitopes (Figure S3a). We identified 3780 MEIS1/2 binding sites, of which 16% were located within 5 kb of a transcription start site (TSS; Figure 3a). 20% of the biding sites overlapped with developmental enhancers linked to putative target genes ((Gorkin et al., 2020), Table S5). Our data predict that MEIS1/2 directly regulates 1218 target genes, either by binding to their TSS or distal enhancers. Many of them (16%) overlapped with genes that were up-regulated in gMeis2-tCROP-seq positive PNs cells (Figure 3b, Table S4-5). *De-novo* motif analysis revealed the previously described MEIS1/2 core hexameric and decameric binding motifs TGACAG and TGATTGACAG, which were highly enriched at the centers of the peaks. These motifs correspond to either the binding of the MEIS homodimer, or the MEIS/PBX heterodimer, respectively (Chang et al., 1997; Shen et al., 1997) (Figure 3c, S3b).

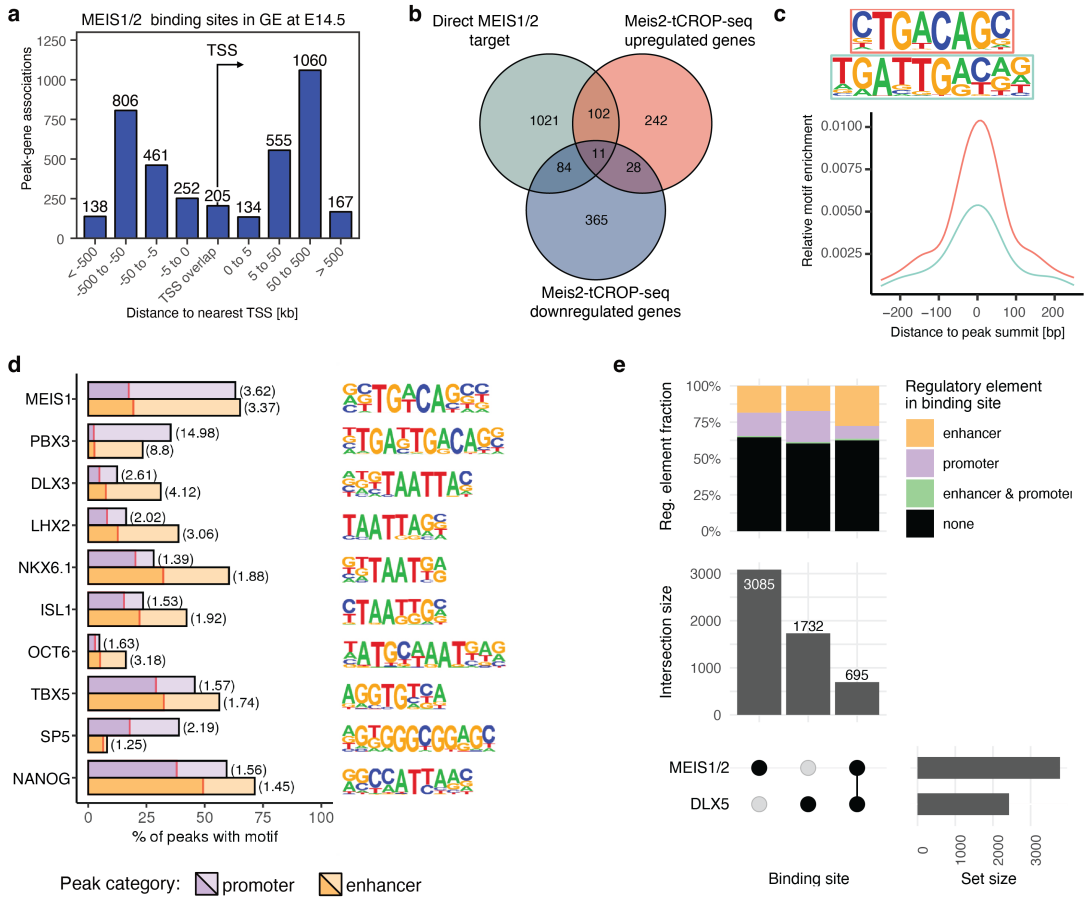


Figure 3: DNA binding of MEIS1/2 in the E14.5 GE. a, Distribution of MEIS1/2-ChIP-seq peaks relative to the nearest transcriptional start site (TSS). **b**, Venn diagram showing overlap between MEIS1/2 target genes and genes up- or downregulated in inhibitory neurons of gMeis2-tCROP-seq. **c**, *De novo* identified MEIS1/2 binding motifs and their position relative to peak summits. **d**, Motif occurrence of selected known motifs enriched within enhancer- or promoter-overlapping MEIS1/2 binding sites (light bars) compared to G/C-matched reference sequences (dark bars), with fold-enrichment in parentheses. **e**, Overlap between binding sites of MEIS1/2 and DLX5 (bottom), with respective distribution of binding sites overlapping promoter and/or enhancer regions.

Binding motifs containing the core sequence TAATT were strongly enriched in MEIS1/2 ChIP-seq peaks, and enriched at enhancers compared to TSS-associated regions. This motif is shared by several homeodomain TF families including those of DLX, LHX and ISL (Figure 3d) (Leung et al., 2022), of which several members are expressed in the GE (Mayer et al., 2018; Leung et al., 2022; Flames et al., 2007). Among them, we found the strongest enrichment for the binding motif of DLX3 (Figure 3d).

All DLX transcription factors share a common conserved motif, of which DLX1, DLX2, DLX5, and DLX6 are known to be master regulators of inhibitory neuron development in the forebrain (Lindtner et al., 2019; Panganiban and Rubenstein, 2002). MEIS and DLX TFs have been shown to interact with each other directly in the olfactory bulb (Agoston et al., 2014) and in a cell culture system (Jolma et al., 2015). Because *Meis2* and *Dlx5* are co-expressed in PN precursor cells of the GE (Figure S4g, S6a), we next tested if MEIS2 and DLX5 interact in the GE. First, we compared the binding sites of MEIS1/2 with those of a recently published DLX5 ChIP-seq dataset in mouse GE (Lindtner et al., 2019). Numerous MEIS1/2 binding sites (695; 18%) overlapped with DLX5 binding sites. Remarkably, the proportion of enhancers at shared (MEIS1/2-DLX5) binding sites was significantly increased compared to MEIS1/2- and DLX5-exclusive binding sites (Figure 3e; $p = 3.694e\text{-}16$, Chi^2-test). The spacing and orientation of MEIS and DLX motifs have previously been described *in vitro*, and changes in spacing between co-transcription factors have been shown to affect gene regulatory capacity (Jolma et al., 2015; Ng et al., 2014; Jindal and Farley, 2021). In our data, the most common motif spacing was 2-4 bp. In contrast to published *in vitro* experiments that observed a fixed spacing of 2 bp between MEIS1 and DLX3 (Jolma et al., 2015), we observed a wider range of spacing (Figure S3c). Together, our findings suggest a potential cooperative role of MEIS1/2 and DLX5 in the fate determination of GE-derived neurons.

Functional link between MEIS2/DLX5 and PN fate

To investigate the possibility of a functional link between MEIS2 and DLX5 during PN development, we performed a series of dual luciferase reporter assays to measure the activity of select genomic

enhancers in the presence of MEIS2, DLX5, or both. To select enhancers with activity in the developing forebrain, we intersected MEIS1/2-DLX5 co-binding sites from ChIP-seq data with the VISTA *in vivo* enhancer database (Visel et al., 2007) (Figure S3d). Additionally, we confirmed the accessibility of respective genomic regions utilizing published scATAC-seq data of the LGE and MGE (Rhodes et al., 2022) (Figure 4a). We chose two enhancers (hs1080 and hs956) of the transcription factor *Foxp2*, which both contained MEIS/DLX motifs with a spacing of 3 bps (Visel et al., 2007; Visel et al., 2013) (Figure 4a, Figure S4a-c,e). *Foxp2* is expressed in precursors of GABAergic PNs (Figure S4g), has previously been implicated in PN development (den Hoed et al., 2021; French and Fisher, 2014), and is one of the genes that we found to be downregulated in gMeis2 tCROP-seq experiments (Table S4). We co-transfected Neuro2a cells encoding either *Dlx5*, *Meis2*, or both with a plasmid containing an enhancer upstream of a minimal promoter. If the enhancer can be activated by DLX5 or MEIS2, the transfected cells would produce measurable luciferase activity. Both *Foxp2* enhancers were modestly activated in the presence of DLX5, whereas MEIS2 alone did not significantly activate either enhancer (Figure 4b-c). Strikingly, MEIS2 and DLX5 together strongly potentiated the DLX5-induced activation of the *Foxp2* enhancers. As expected, PBX1, a known interaction partner of MEIS2 (Hyman-Walsh et al., 2010), increased the effect of MEIS2 (Figure S4d,f). These results suggest that MEIS2 and DLX5 bind cooperatively at specific binding sites of enhancers to regulate *Foxp2* expression.

Mutations in the *MEIS2* gene have been linked to intellectual disability, cardiac defects and facial phenotypes (Louw et al., 2015; Verheije et al., 2019; Giliberti et al., 2020; Gangfuß et al., 2021). At least four patients with severe disease carry either a frameshift mutation, an in-frame deletion or a missense mutation of a single highly conserved amino acid (Arg333) located in the MEIS2 homeodomain (Giliberti et al., 2020; Gangfuß et al., 2021). We tested whether the p.Arg333Lys missense variant (MEIS2*333) can activate the *Foxp2* enhancer hs956. DLX5-dependent joint activation of hs956 was greatly reduced with MEIS2*333 compared to wild-type MEIS2 (Figure 4e).

We next investigated whether the cooperation of MEIS2 and DLX5 at co-binding sites also

activates a putative regulatory enhancer (enhD1) of *Drd1* (Gorkin et al., 2020). *Drd1* encodes for the dopamine receptor D1, which is a top marker of D1-type medium spiny projection neurons (D1-MSN; PN:Foxp1/Isl1, PN:Isl1/Bcl11b, PN:Ebf1/Zfp503) in the striatum (Gerfen and Surmeier, 2011) (Figure 4d, Figure S4g). Its gene expression was strongly reduced in PN clusters in gMeis2 tCROP-seq experiments (Table S4, Figure 2f). EnhD1 is located in a topologically associated domain (TAD) with the *Drd1* gene body (Figure S5a)(Bonev et al., 2017). Furthermore, enhD1 contained pronounced ChIP-seq peaks (Figure 4d), and multiple MEIS/DLX co-binding motifs (Figure S5b). MEIS2 did not activate enhD1, but it potentiated the effect of DLX5 on the activity of enhD1, in a concentration-dependent manner (Figure 4e-f). Furthermore, the cooperative activation of enhD1 by MEIS2 and DLX5 was greatly reduced with the mutated version of MEIS2 (Meis2*333; Figure 4f). A truncated version of enhD1 in which a portion (TG) of the MEIS binding motif was removed at multiple sites (Figure S5b) showed reduced activation by MEIS2/DLX5 compared with the unmodified truncated enhD1 (Figure 4g). Taken together, this suggests that the cooperation of MEIS2 and DLX5 at specific co-binding sites within regulatory elements activates gene expression of TFs that promote PN fate.

Next, we tested whether MEIS2 is able to activate the promoters of *Pbx3*, *Tshz1* and *Six3*. All three genes are marker genes for different PN clusters, and they all contain binding sites for MEIS in their promoters (Figure 1c). MEIS2 did not activate these promoters (Figure S4h). Interestingly, even the *Tshz1* promoter, which contains both DLX5 and MEIS1/2 motifs, was not activated by MEIS2, nor was MEIS2 able to enhance the DLX5-induced activation of this promoter (Figure 4h-i). This may be because the motifs for DLX and MEIS1/2 are far apart at this promoter.

Our data suggest that in the GE, MEIS2 requires the presence of DLX5 to bind and co-activate enhancers with specific co-binding sites. This process induces gene expression related to PN development. We tested a total of ten enhancers or promoters of genes which are known to be important for inhibitory neuron development using a dual-luciferase reporter assay, and the results support this model: Of the enhancers tested, only the LGE-specific enhancer of *Aldh1a3*, enhAldh1a3, which lacks a MEIS1/2-DLX5 co-binding site, was strongly activated by MEIS2

alone (Figure 4j-k; Figure S4g-h). *Aldh1a3* encodes an enzyme that synthesizes retinoic acid in LGE precursors at E12.5 (Molotkova et al., 2007; Toresson et al., 1999) and is essential for the differentiation of striatal PNs (Chatzi et al., 2011). *Aldh1a3* was greatly downregulated in several clusters in the gMeis2 tCROP-seq experiments (Table S4). It remains unclear whether MEIS2 is able to activate enhAldh1a3 on its own, or whether another cofactor present in Neuro2a cells is required.

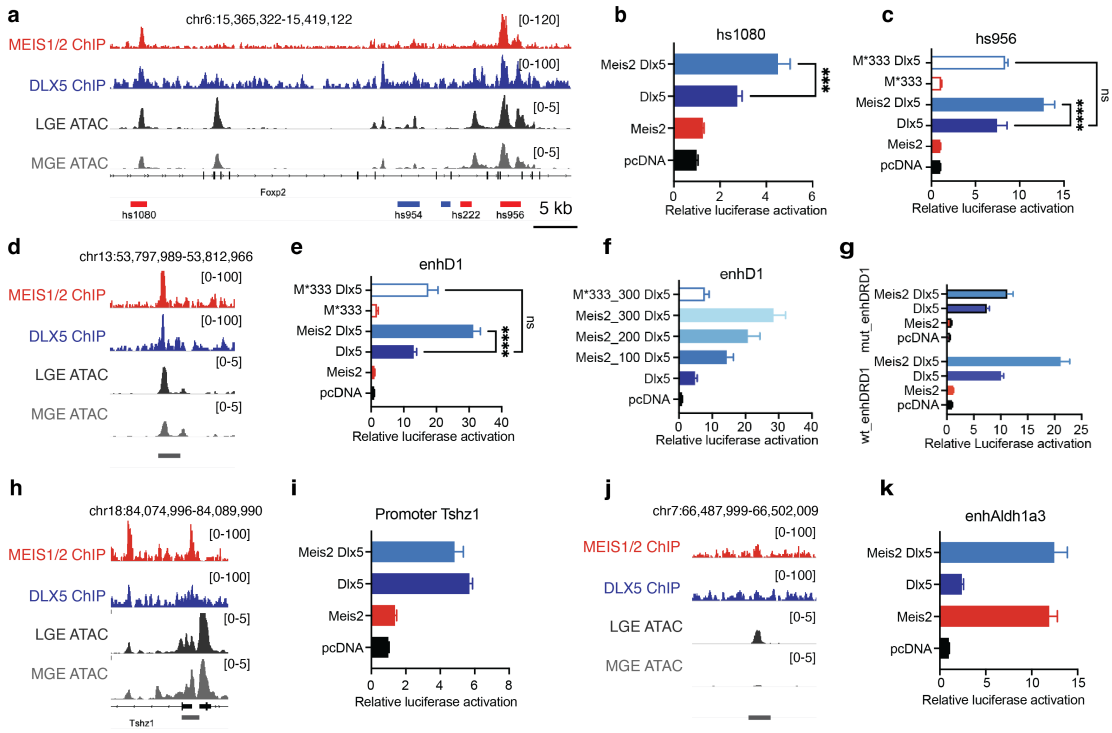


Figure 4: Cooperation between MEIS2 and DLX5 to activate projection neuron regulatory elements. a, Representative profiles of MEIS1/2 (red) and DLX5 (blue) ChIP-Seq at E14.5 and E13.5 respectively, as well as scATAC-seq from LGE (dark grey) and MGE (grey) at E12.5 are shown at the *Foxp2* gene locus. DLX5 ChIP-Seq data from (Lindtner et al., 2019); scATAC-seq data from (Rhodes et al., 2022). **b**, Luciferase activity driven by the enhancer hs1080, co-transfected with Meis2 and Dlx5 expression vectors in Neuro2 cells. **c**, Luciferase reporter assays of the enhancer hs956. **d**, Representative profiles of the *Drd1* gene enhancer enhD1. **e**, Luciferase reporter assays of enhD1. **f**, Luciferase reporter assays of enhD1, co-transfected with Dlx5 and increasing concentration of Meis2, or with Meis2*333. **g**, Luciferase reporter assays of the wild-type or mutated, shorter version of enhD1. **h**, Representative profiles of the *Tshz1* promoter. **i**, Luciferase reporter assays of the *Tshz1* promoter. **j**, Representative profiles of the *Aldh1a3* enhancer enhAldh1a3. **k**, Luciferase reporter assays of enhAldh1a3. In panels b, c, e, f, g, i, and k, individual data points are shown with bars representing the mean ± s.e.m. from three independent experiments, each performed in triplicate. One-way ANOVA test: *P <0.05, **P <0.01, ***P <0.001 and ****P <0.0001.

Spatial patterning and the functional activity of Meis2 in the GE

PNs of the striatum originate largely in the LGE, and many IN types, e.g., those of the cortex, originate in the MGE and CGE (Knowles et al., 2021; Lim et al., 2018; Bandler et al., 2017). MEIS2 is initially expressed broadly in the ventricular zone (VZ) of the LGE, CGE and MGE. In neuronal precursors of the subventricular (SVZ) and mantle zones (MZ), a spatial pattern of MEIS2 expression emerges, where MEIS2 continues to be highly expressed in the LGE, but is absent in the MGE (Figure S6a) (Toresson et al., 1999; Su et al., 2022). We next asked how the function of MEIS2, as a DLX-dependent activator of PN development, acquires LGE selectivity. First, we tested if LHX6, which is exclusively expressed in the MGE and enriched in the SVZ and MZ, and whose activity is required for the specification of cortical IN subtypes (Figure S6a) (Sandberg et al., 2016; Zhang et al., 2013; Cesario et al., 2015), affects the function and expression of MEIS2. We intersected ChIP-seq peaks in the GE of MEIS1/2, DLX5 (Lindtner et al., 2019) and LHX6 (Sandberg et al., 2016) (Figure S3e, Table S5). Out of 151 MEIS1/2-DLX5-LHX6 overlapping peaks, 41 were within Vista enhancers, and 28 of these enhancers showed activity in the developing forebrain (Figure S6b-c). We selected three of them to perform luciferase reporter assays (Figure 5a-f): (1) hs1041, an enhancer of the *Tle4*, which encodes transcription co-repressor 4, (2) hs956, an enhancer of *Foxp2*, and (3) hs748, an enhancer of *Zfp503*, which encodes the zinc finger protein TF 503 (NOLZ1). Genes regulated by the selected enhancers are known to play a role in striatal development (Shang et al., 2022; den Hoed et al., 2021; French and Fisher, 2014; Su-Feher et al., 2021), were expressed in PN precursors (Table S2), and were reduced in several clusters in the gMeis2 tCROP-seq experiments (Table S4). Consistent with the above findings, MEIS2 strongly potentiated the DLX5-mediated activation of hs1041, hs956, and hs748 reporters. LHX6 alone had little to no effect on the activation of these enhancers. However, co-expression of LHX6 with MEIS2 and DLX5, resulted in strong suppression of enhancer activity in all three cases (Figure 5g-i). This suggests that LHX6, whose expression is spatially restricted to the MGE, suppresses the DLX5-MEIS2-induced enhancer activation. This demonstrates how, in the unfolding

of developmental genetic programs in the GE, spatial factors are translated into selective enhancer activation. To gather further evidence for this mechanism, we screened 15 Vista enhancers with overlapping ChIP-seq peaks for LHX6, MEIS1/2 and DLX5 (Figure S6c. None of them exhibited robust activity in the MGE, supporting our interpretation that LHX6 prevents the activator function of MEIS2 in the MGE.

Next, we explored the putative enhancer of *Meis2*, enhMeis2 (Gorkin et al., 2020), which also contained MEIS1/2-DLX5-LHX6 co-binding sites (Figure 5j). MEIS2 strongly potentiated the DLX5-mediated activation of enhMeis2 (Figure 5k), suggesting that in the presence of DLX5, MEIS2 can promote its own expression via the activation of enhMeis2. Self-activation has already been reported previously for *Meis* genes (Bridoux et al., 2020). Strikingly, LHX6 strongly repressed the MEIS2-DLX5 mediated activation of enhMeis2, suggesting that LHX6 suppresses the expression of MEIS2, consistent with a recent *Lhx6* knockout study in mice (Asgarian et al., 2022). This may explain the absence of MEIS2 in the SVZ and MZ of the MGE, and adds another level of regulation aimed at suppressing PN fate in MGE precursors (Figure 5l, S8a).

Meis2 and Lhx6 alter gene modules in PNs and INs

To explore how the depletion of embryonic TFs alters postnatal cell-type composition and identity, we performed pooled tCROP-seq experiments with sgRNAs for *Meis2* (gMeis2), *Lhx6* (gLhx6), *Tcf4* (gTcf4), and LacZ (gLacZ, control). Like LHX6, TCF4 is enriched in INs, but is expressed in all GEs (Kim et al., 2020) (Figure S6a). A mixture of plasmids encoding one of the sgRNAs (Table S1), CAS9-EGFP (a gift from Randall Platt) ((Platt et al., 2014)), and a pB-helper plasmid was targeted to the GE via *in utero* electroporation at E12.5 (Figure 6a-b). At P7 we collected 35 pups, enriched tdTomato/eGFP positive cells with FACS, and performed pooled scRNA-seq. Ten scRNA-seq datasets were combined in silico, clustered, and annotated based on known marker genes (Figure 6c-d, S7a, Table S6). All three perturbations had a significant effect on the composition of cell types compared to the gLacZ control (Figure 6e-f). Cells expressing gLhx6 showed an increased proportion of medium spiny projection neurons (D1/D2 MSNs), olfactory bulb (OB) precursors, and

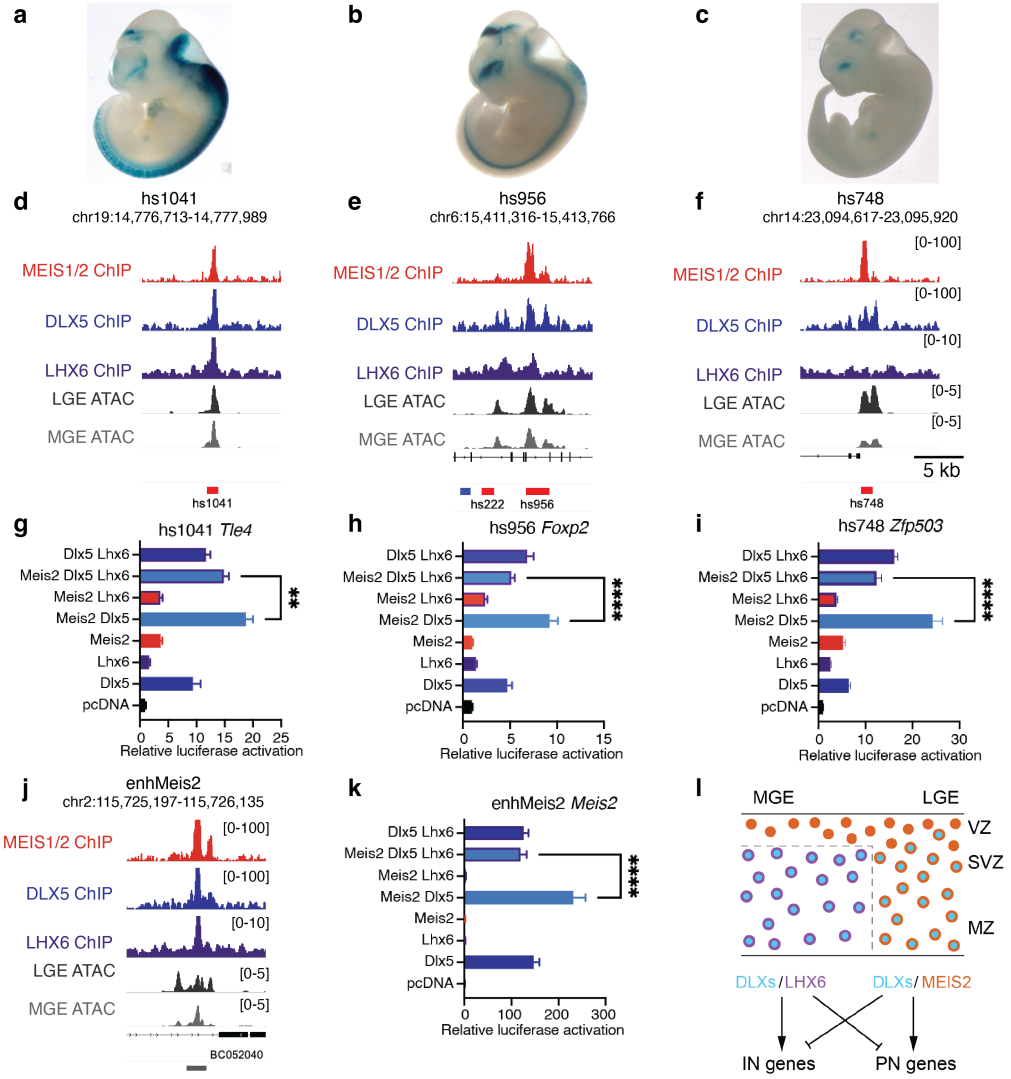


Figure 5: Regulation of LGE enhancers by MEIS2, DLX5 and LHX6. **a**, **b**, **c**, hs1041, hs956, hs748 enhancers drives LacZ expression at E12.5 (Visel et al., 2007). **d**, **e**, **f**, Representative tracks of GE ChIP-Seq of MEIS1/2 at E14.5 (red), DLX5 at E13.5 (blue) (Lindtner et al., 2019), LHX6 at E13.5 (purple) (Sandberg et al., 2016) and scATAC-seq (Rhodes et al., 2022) from the LGE (dark grey) and MGE (grey) at E12.5. **g**, **h**, **i**, Luciferase activity driven by hs1041, hs956, and hs748 enhancers co-transfected with Meis2, Dlx5, and Lhx6 expression vectors in Neuro2a cells. **j**, Representative tracks of *Meis2* gene enhancer enhMeis2. **k**, Luciferase reporter assays of enhMeis2. **l**, A model of the proposed actions of MEIS2, DLX5, and LHX6. MEIS2 promotes projection neuron fate in the presence of DLX proteins, while LHX6 suppresses it and promotes interneuron fate. MGE, medial GE; LGE, lateral GE; SVZ, subventricular zone; VZ, ventricular zone; MZ, mantle zone. In g, h, i, f, and k, individual data points are shown with bars representing mean ± s.e.m from three independent experiments, each performed in triplicate. One-way ANOVA test: *P<0.05, **P<0.01, ***P<0.001 and P****<0.0001.

INs compared to gLacZ. An increase of CGE INs after *Lhx6* deletion has previously been reported (Vogt et al., 2014). Consistent with our embryonic tCROP-seq data, the proportion of INs was also increased in gMeis2. In cells expressing gMeis2, amygdala intercalating cells (ITCs) were depleted and the number of olfactory bulb inhibitory neurons and oligodendrocyte progenitor cells (OPCs) was reduced (Figure 6e-f). The ITC depletion is consistent with the E16.5 tCROP-seq results, where the PN:Tshz1/Pbx3 cluster (likely corresponding to immature ITC cells (Kuerbitz et al., 2018)) had the most DEGs (Figure 1f). gTcf4 expression had a more modest effect on cell proportions, showing only a slight reduction in inhibitory neurons in the olfactory bulb. Furthermore, both gMeis2 and gTcf4 showed a reduced number of astrocytes. Across all clusters, gLhx6, gMeis2, and gTcf4 positive cells had a total of 90, 58, and 7 DEGs respectively (Figure 6g-h, Table S7). Many of them were marker genes specifically expressed in IN or PN cell types (Table S6-7). gLhx6 perturbed cells were enriched for PN specific genes (*Isl1*, *Foxp1*, *Ebf1*, *Adora2a*, *Drd1*, *Six3*). By contrast, gMeis2 DEGs were enriched for IN-specific genes (*Maf* and *Prox1os*) and depleted for PN-specific genes (*Mpped2* and *Pbx3*). Our data support the idea that MEIS2 primarily induces PN fate and LHX6 primarily induces IN fate. scRNA-seq data are highly heterogeneous and have numerous zero counts, making it challenging to detect subtler perturbation-based biological changes in single cell datasets. To overcome these limitations, we utilized Hotspot (DeTomaso and Yosef, 2021), a tool that identifies co-varying groups of genes (modules). Each cell was assigned a gene module score, with higher scores indicating higher association with that module. We identified 8 Hotspot gene modules (Figure S7b), 4 of which were neuronal (Figure 6i, S7c). Module 5 is represented mostly in olfactory bulb neuroblasts and contains genes enriched for neuronal differentiation. Module 4 is represented in MSN cell types and contained MSN marker genes (e.g., *Foxp1*) and genes involved in retinoic acid receptor signalling (*Rarb*, *Rxrg*). The retinoic acid pathway is involved in the switch between proliferation and differentiation (Berenguer and Duester, 2022), which is essential for striatal development (Chatzi et al., 2011). Module 8 was represented in OB precursors and ITC cells. This module contained *Meis2* as well as some of its target genes, such as *Pbx3* and *Etv1* (Table S4). Module 6 was represented in the OB-Cpne4 population and was characterized by genes

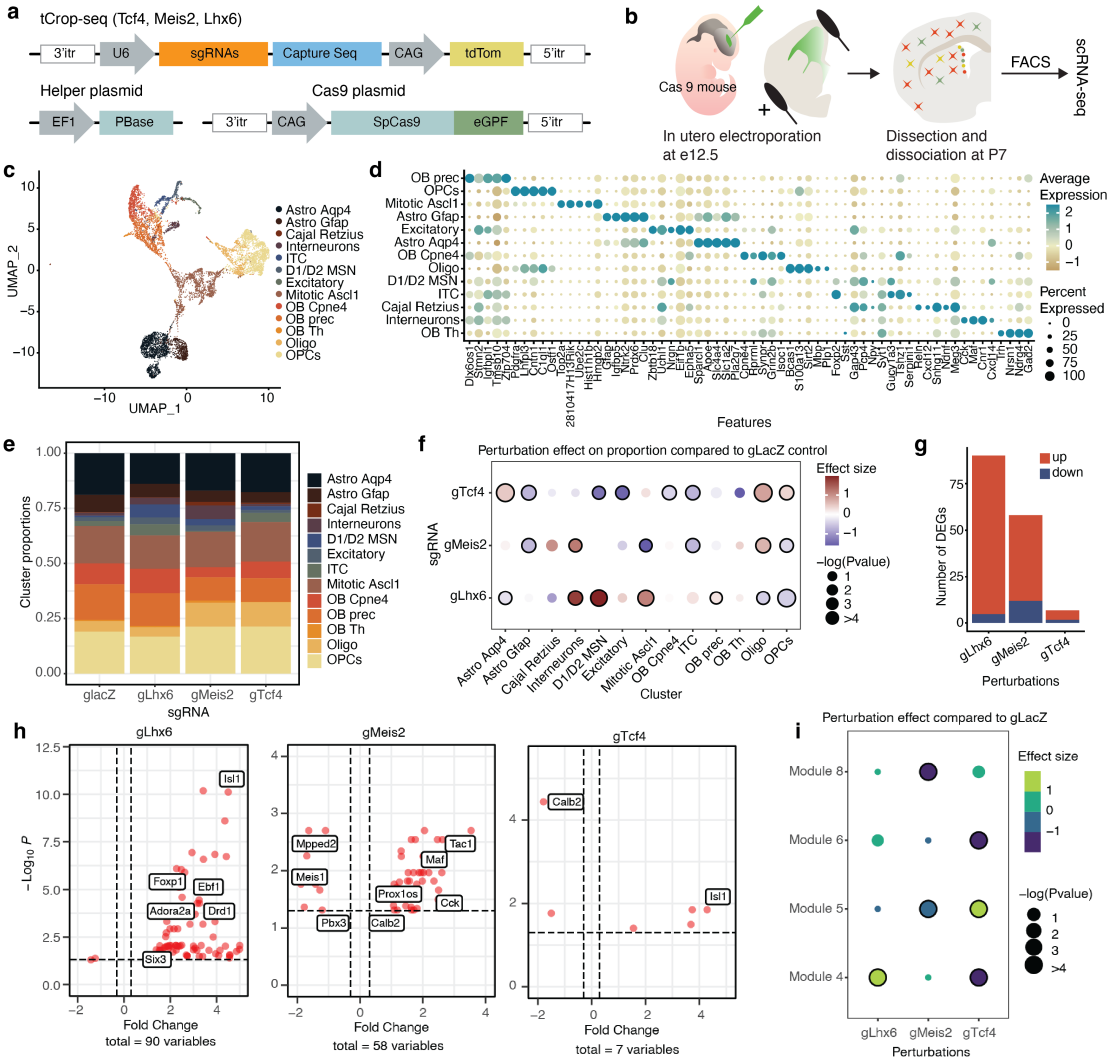


Figure 6: Embryonic disruption of transcription factors alters postnatal cell types a, Vector maps of tCrop-seq. **b**, Schematic illustrating the workflow of tCrop-seq. **c**, UMAP plot of the P7 data colored by cell type. **d**, Dot plot showing the top 5 marker genes of each cell type. OB, olfactory bulb cells; OPC, oligodendrocyte progenitor cells; ITC, intercalated cells; MSN, medium spiny neurons; Oligo, oligodendrocyte, Astro, astrocytes. **e**, Cell type compositions for each sgRNA. **f**, Dot plot showing the effect of perturbation on cell type composition for each sgRNA compared to the control (glacZ). Black outline indicates statistical significance (p-val < 0.05). **g**, Bar plot showing the number of differentially expressed genes detected in each sgRNA. **h**, Differential gene expression analysis of inhibitory neurons in each sgRNA compared to control (glacZ). Volcano plot showing differentially expressed genes that meet the cutoff criteria (FDR < 0.05, avg_logFC > 0.5). **i**, Dot plot showing the effect of perturbation by sgRNAs on the module scores of inhibitory modules.

involved in calcium response and synapse organization. We fitted a linear regression model that accounted for the batch and number of genes, and extracted the effect sizes to estimate how the module scores in the perturbed cells deviated from gLacZ control cells (Jin et al., 2020). For the three TFs, the perturbations had significant effects across different modules (FDR-corrected $P < 0.05$; Figure 6i). The perturbation of *Lhx6* was positively associated with the expression of module 4, consistent with the change in cell proportion and change in differentially expressed genes. The perturbation of *Meis2* lowered the expression of both modules 8 and 5. The perturbation of *Tcf4* had a significant effect across modules 6, 5, and 4, consistent with previous findings showing that TCF4 is a key facilitator of neurogenesis and neuronal differentiation (Figure 6i) (Mesman et al., 2020; Teixeira et al., 2021). Taken together, the tCROP-seq data at P7 indicate a marked influence of MEIS2, LHX6, and TCF4 on PN and IN specification.

Discussion

Our findings contribute to an overall picture in which spatial selective enhancer activation, rather than hierarchical TF cascades, is instrumental in the early imprinting of GABAergic identities (Figure S8). Different GABAergic cell types arise from regional differences in the specification of GE progenitors, which are initially established by morphogenetic molecules such as retinoic acid (RA, LGE) (Chatzi et al., 2011), fibroblast growth factor (FGF) 8 and sonic hedgehog (SHH, MGE) (Storm et al., 2006; Molotkova et al., 2007), FGF12 and FGF15 (CGE) (Borello et al., 2008; Shohayeb et al., 2021), and their downstream TFs, such as MEIS2 (LGE), NKX2.1 & LHX6 (MGE), and NR2F1/2 (CGE). Our results depict how these spatial factors are utilized downstream for selective enhancer activation.

We show, in agreement with Su et al., 2022 (Su et al., 2022), that MEIS2 is required for the activation of PN-specific genes and promotes LGE PN fate. Furthermore, we demonstrate that MEIS2 requires the presence of DLX5 to accomplish this, and that together MEIS2 and DLX5 activate numerous regulatory elements containing specific co-binding sites. The tissue specificity of

DLX5, and likely other members of the DLX family in the GE, directs the functional activity of MEIS2 to regulatory sites related to GABAergic PN development. This is consistent with a proposed model of TALE transcription factors (e.g. MEIS) acting as broad co-activators of homeobox genes (Bridoux et al., 2020). Multiple studies have demonstrated that MEIS proteins interact with other transcription factors, such as PBX, HOX, TBX, and PAX6, to promote differentiation in the limbs, heart, lens, hindbrain, and olfactory bulb (Schulte and Geerts, 2019; Bridoux et al., 2020; Delgado et al., 2021; Selleri et al., 2019; Agoston et al., 2014). MEIS2 appears to act in a highly context-dependent manner. For example, MEIS1/2 ChIP-seq from the retina (Dupacova et al., 2021) shows little overlap with ChIP-seq from the GE. Direct interaction between MEIS1/2 and DLX transcription factors, as well as co-regulation of gene expression, has been reported in the olfactory bulb and in *in vitro* experiments (Agoston et al., 2014; Jolma et al., 2015).

We provide evidence that the co-activation of select enhancers by MEIS2/DLX5 is absent in the MGE. This spatial component within the GE is partially effectuated by LHX6, whose expression is likely restricted to the MGE through FGF8/SHH/NKX2.1 signaling (Storm et al., 2006; Molotkova et al., 2007; Sandberg et al., 2016). We show that LHX6 suppresses the activation induced by MEIS2/DLX5 via specific binding sites close to the MEIS1/2-DLX5 co-binding sites. The suppression by LHX6 could be mediated via competition of LHX6 with DLX for the common DNA binding motif TAATT (Sandberg et al., 2016). Alternatively, LHX6 could restrict the interaction of MEIS2/DLX5 with DNA through direct binding to DLX5 or MEIS2. LHX6 belongs to the LIM domain homeodomain (LIM-HD) protein family, which is characterized by two cysteine-rich LIM domains for protein-protein interactions and a homeodomain for binding DNA (Hobert and Westphal, 2000). For example, LHX6 directly interacts with PITX2 to inhibit its transcriptional activities (Zhang et al., 2013).

We show that LHX6 suppresses regulatory elements of *Meis2*, likely resulting in repression of *Meis2* gene expression in the SVZ and MZ of the MGE. Consistently, *Meis2*, as well as the PN marker genes *Pbx3* and *Foxp1*, have been shown to be up-regulated in E14.5 Lhx6 knockout cells collected from the cortex (Asgarian et al., 2022). Furthermore, conditional knockout of NKX2-1,

which acts upstream of LHX6, has been shown to result in increased transcription of MEIS2 in the SVZ of the MGE (Sandberg et al., 2016) and an enrichment of repressive regulatory elements in motifs consistent with the binding site of MEIS2 (Sandberg et al., 2018).

MEIS2 could inhibit IN fate through the activation of repressive transcription factors such as ISL1, FOXP1/2, and SIX3, via co-repressors such as TLE1/4, or by promoting the expression of microRNAs (miRNAs). We identified several miRNA host genes that were downregulated in Meis2-tCROP-seq: *Mir124-2hg Gm27032*(miR-124a-2), *Arpp21* (miR-128-2), and *Gm27032* (miR-124a-3; Table S4). miR-124 and miR-128 are some of the most abundant and highest enriched miRNAs in the adult mouse and human brains (Zolboot et al., 2021). miR-128 deficiency in D1-MSNs leads to juvenile hyperactivity, followed by lethal seizures at 5 months of age (Tan et al., 2013).

Haploinsufficiency of the transcription factor MEIS2 in humans results in an autosomal dominant disease characterized by multiple congenital malformations, mild-to-severe intellectual disability with poor speech, and delayed psychomotor development (Louw et al., 2015; Douglas et al., 2018; Giliberti et al., 2020; Gangfuß et al., 2021; Zhang et al., 2021). The amino acid Arg333, located in the homeodomain of MEIS2, is highly conserved across species and isoforms (Longobardi et al., 2014) and was found mutated in at least four patients with severe disease (Giliberti et al., 2020; Gangfuß et al., 2021). Our study found that the missense mutation p.Arg333Lys led to a strong decrease in enhancer activation compared to wild-type MEIS2. Due to the location of Arg333 in the homeodomain of MEIS2, it is likely that mutations in this amino acid interfere with the protein's DNA binding ability. This could result in a change in GABAergic cell type proportions, in particular a reduced number of PNs in the striatum, caused by disturbed fate decisions during embryogenesis, and ultimately elicit the disease phenotype seen in affected individuals. The efficiency with which MEIS2 can co-activate selective enhancers suggests a general strategy for implementing spatial information to generate distinct cellular populations. The ability of MEIS2 to induce context-specific cell types may exemplify how certain subsets of cells in different parts of the body are affected in developmental disorders. Further research is needed to fully comprehend the intricate interactions

between transcription factors and cofactors in the regulation of cell fate decisions during GABAergic neuron development and their potential implications in human disease.

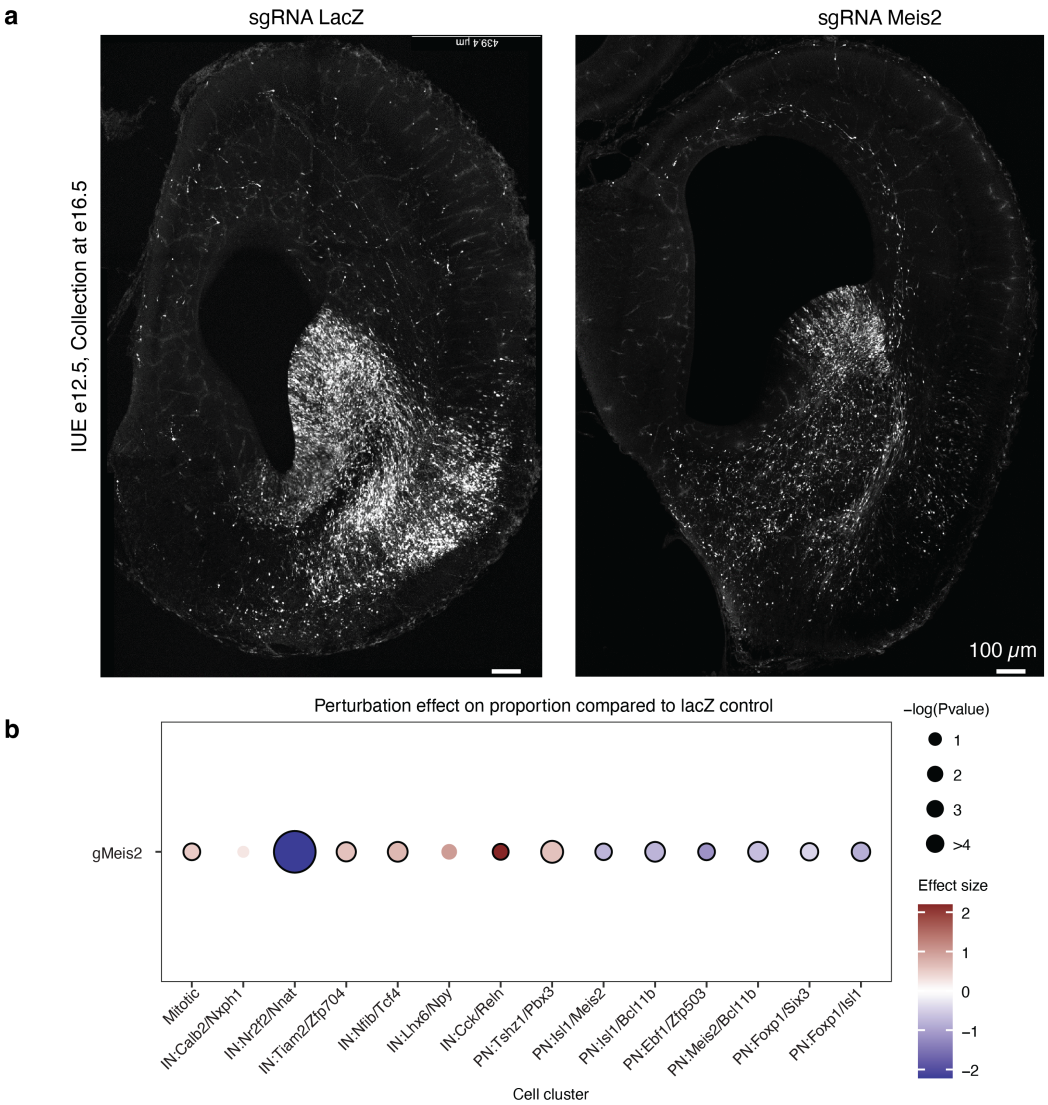
Acknowledgements

We thank members of the Mayer and Winkelmann laboratories for feedback and discussion; J. Kuhl (somedonkey.com) for illustrations; R. H. Kim from the MPIB Next Generation Sequencing core facility, I. Velasques and G. Eckstein from the Genomics Core facility at the Helmholtz Zentrum München (HMGU), M. Spitaler and M. Oster from the MPIB Imaging and FACS core facility, R. Kasper from the MPIBI Imaging core facility and members of the MPIB/MPIBI animal facility for their technical expertise. This work was supported by the Max-Planck Society, the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation program (ERC-2018-STG, grant agreement no. 803984; to C.M.), and the European Commission (SMART GRANT: ERA-NET NEURON, SMART: 01EW1605).

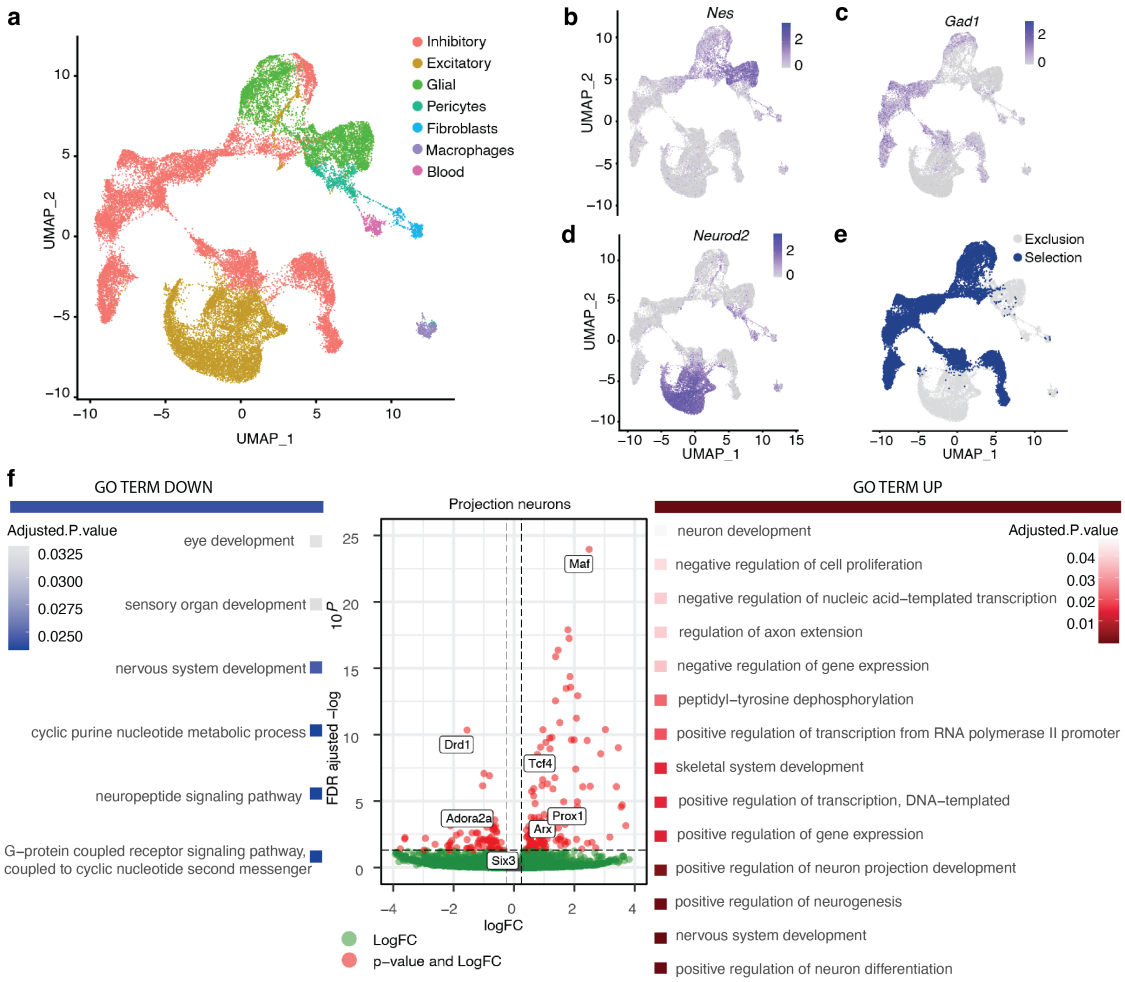
Author contributions statement

E.D. and C.M. conceived the project; M.H. and C.M. developed TrackerSeq; E.D. and C.M. developed tCROP-seq; E.D., I.V. and C.M. conducted the tCROP-seq and TrackerSeq experiments; D.D.L., I.D. and M.T. conducted the MEIS1/2 ChIP-seq experiments; E.D. conducted functional reporter assay experiments; M.H. lead scRNA-seq, tCROP-seq, TrackerSeq computational analyses; V.K. lead ChIP-seq analyses; E.D, M.H., V.K., J.W. and C.M. prepared the manuscript with input from the remaining authors. **Competing interests:** The authors declare no competing interests.

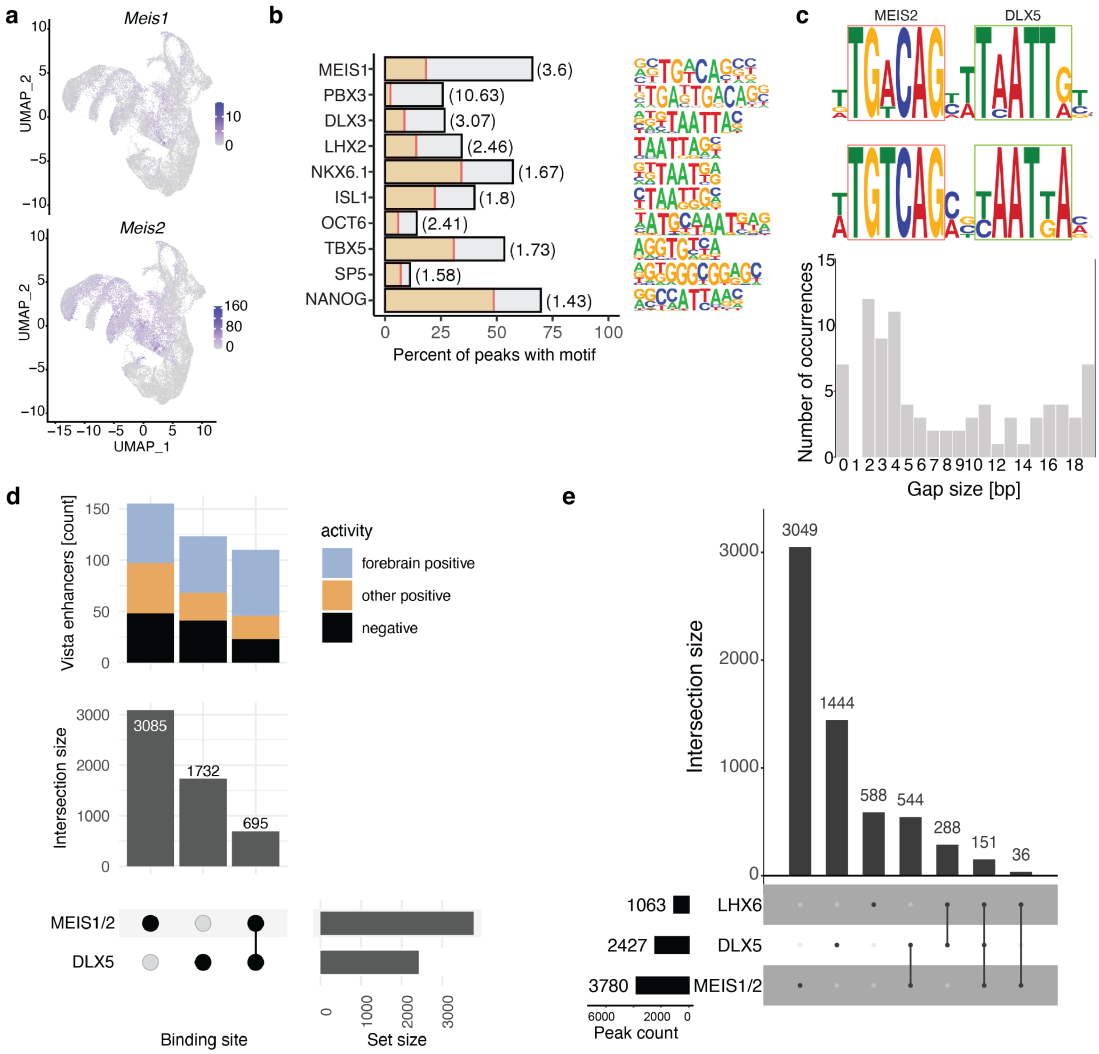
Supplementary Figures



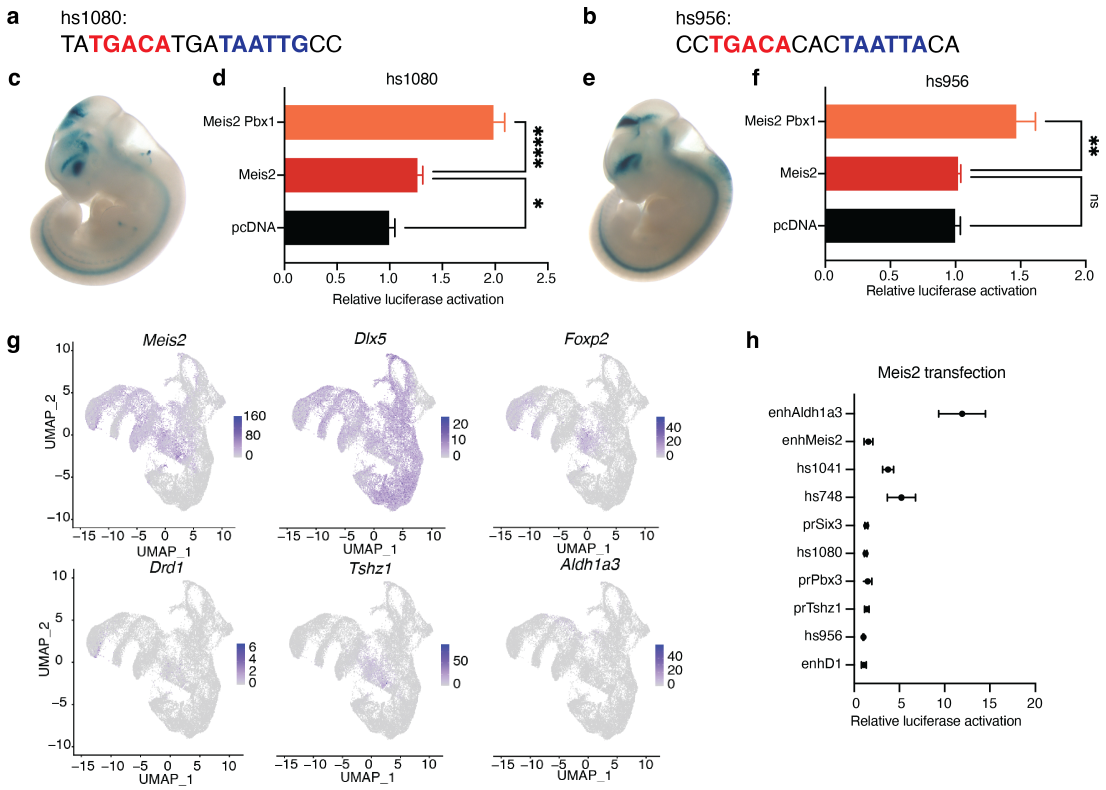
Supplementary Figure 1: In utero electroporation of cells with gLacZ and gMeis2 vectors carrying a TdTomato reporter at E16.5. a, Distribution of cells in the cortex, striatum, and GE. Scale bar, 0.1 mm. **b**, Dot plot showing the effect of perturbation gMeis2 on the cell type proportion of each cell cluster compared to the control (glacZ). Black outline indicates statistical significance (p-val < 0.05).



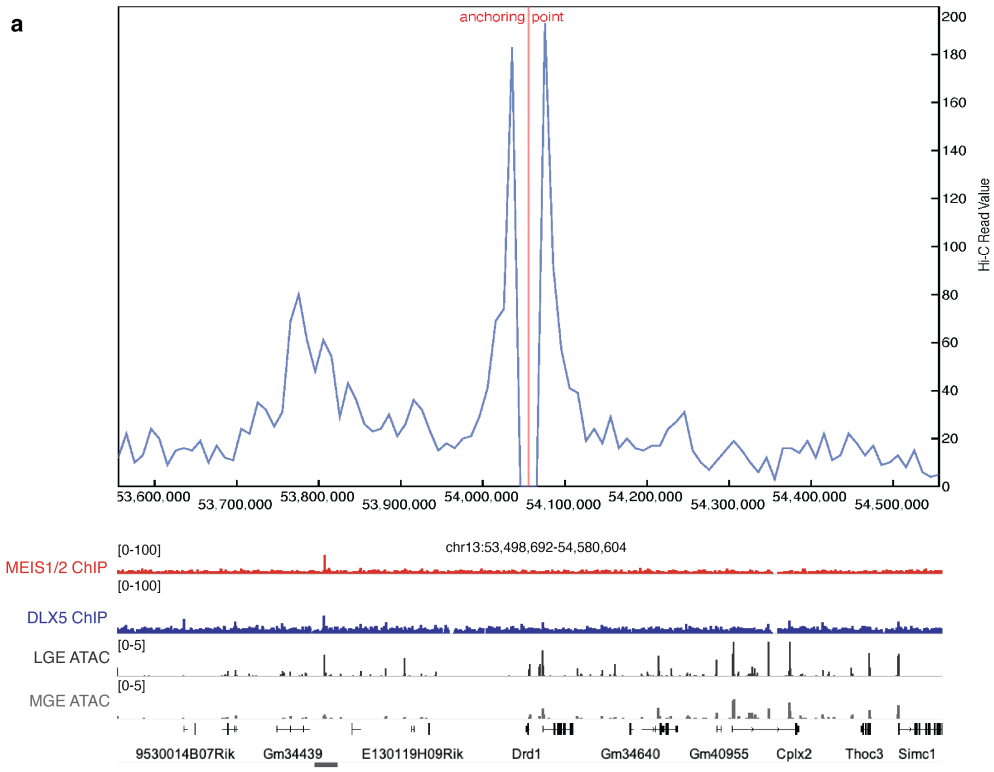
Supplementary Figure 2: Single cell analysis of cell classes at E16. **a**, UMAP plot of E16 data colored by cell classes. **b**, **c**, **d**, Feature plots of the canonical marker genes *Nes*, *Gad1*, and *Neurod2*. **e**, UMAP plot depicting the selection of cells for downstream analysis. **f**, Gene ontology analysis on differentially expressed genes (DEG) of clusters belonging to the projection neuron class.



Supplementary Figure 3: Analysis of MEIS1/2 ChIP-seq. a, Feature plots of *Meis1* and *Meis2* at E16, with expression depicted from gray (low) to purple (high). **b**, Motif occurrence of selected known motifs enriched within all MEIS1/2 binding sites (grey bars) compared to G/C-matched reference sequences (yellow). **c**, Motif spacing analysis of MEIS2 and DLX5 motifs within shared binding sites, with the position weight matrix (PMW) of the most frequent motif configuration on the left and the overall distribution of the DLX5 motif relative to the MEIS2 motif on the right. **d**, Overlap between binding sites of MEIS1/2 and DLX5 (bottom) and their respective distribution within different classes of Vista enhancers (top). **e**, Overlap between binding sites of MEIS1/2, DLX5 and LHX6.

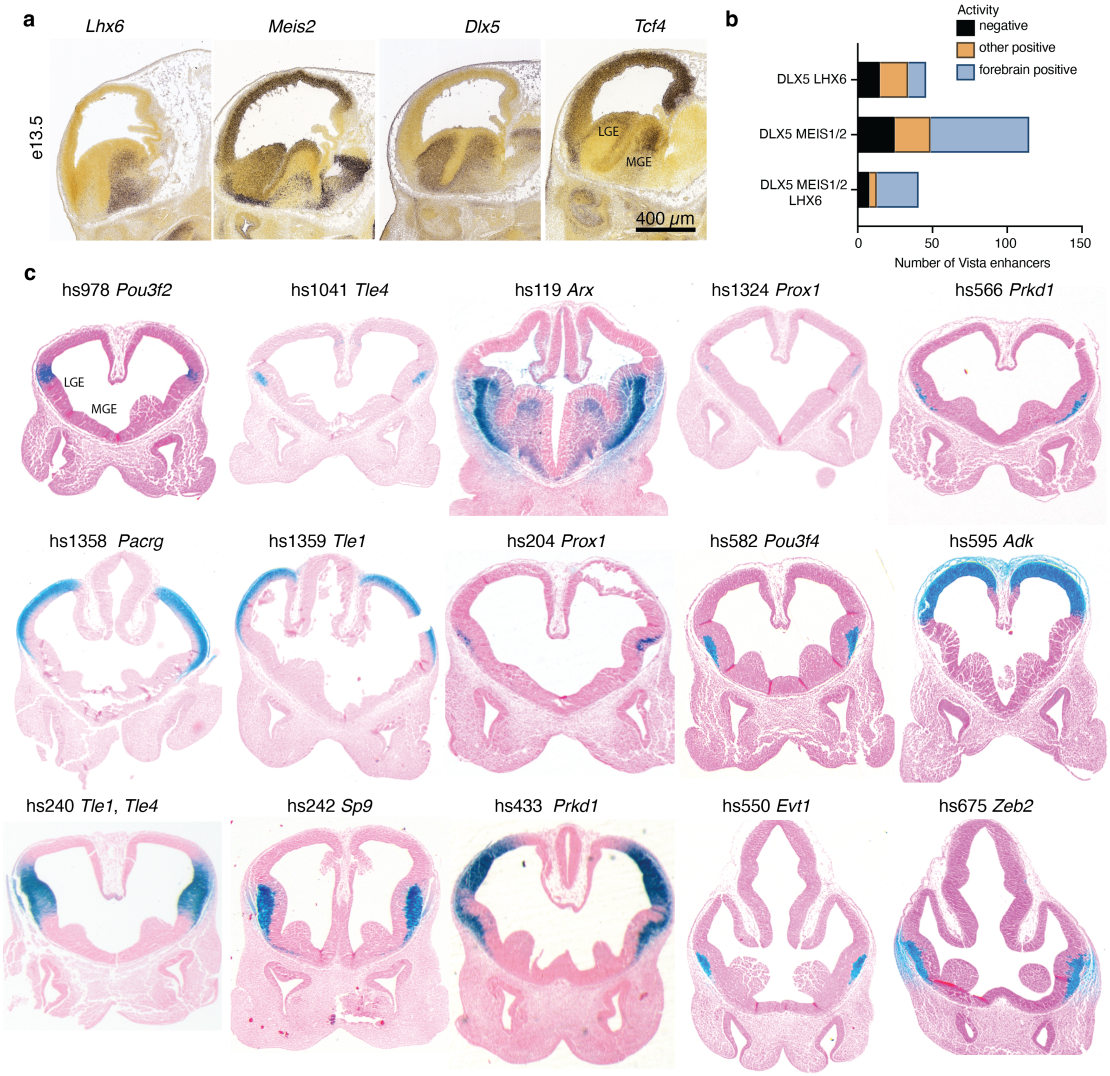


Supplementary Figure 4: Luciferase reporter assay of projection neuron regulatory elements. **a**, **b**, Combined MEIS (red) and DLX (blue) binding motifs found within hs1080 (a) and hs956 (b) enhancers. **c**, hs1080 enhancer drives LacZ expression in E12.5 mouse forebrain (Visel et al., 2007). **d**, Luciferase activity driven by hs1080 enhancer co-transfected with MEIS2 and PBX1 expression vectors in Neuro2 cells. **e**, hs956 enhancer drives LacZ expression in E12.5 mouse forebrain. **f**, Luciferase activity driven by hs956 enhancer. **g**, Feature plots of *Meis2*, *Dlx5*, *Foxp2*, *Drd1*, *Tshz1* and *Aldh1a3* at E16, with expression depicted from gray (low) to purple (high). **h**, Luciferase activity driven by regulatory elements transfected with Meis2 expression vectors in Neuro2 cells (pooled from different experiments). In d, f, and h, individual data points are shown with bars representing mean ± SEM from three independent experiments, each performed in triplicate. One-way ANOVA test: *P <0.05, **P <0.01, ***P <0.001 and ****P <0.0001.

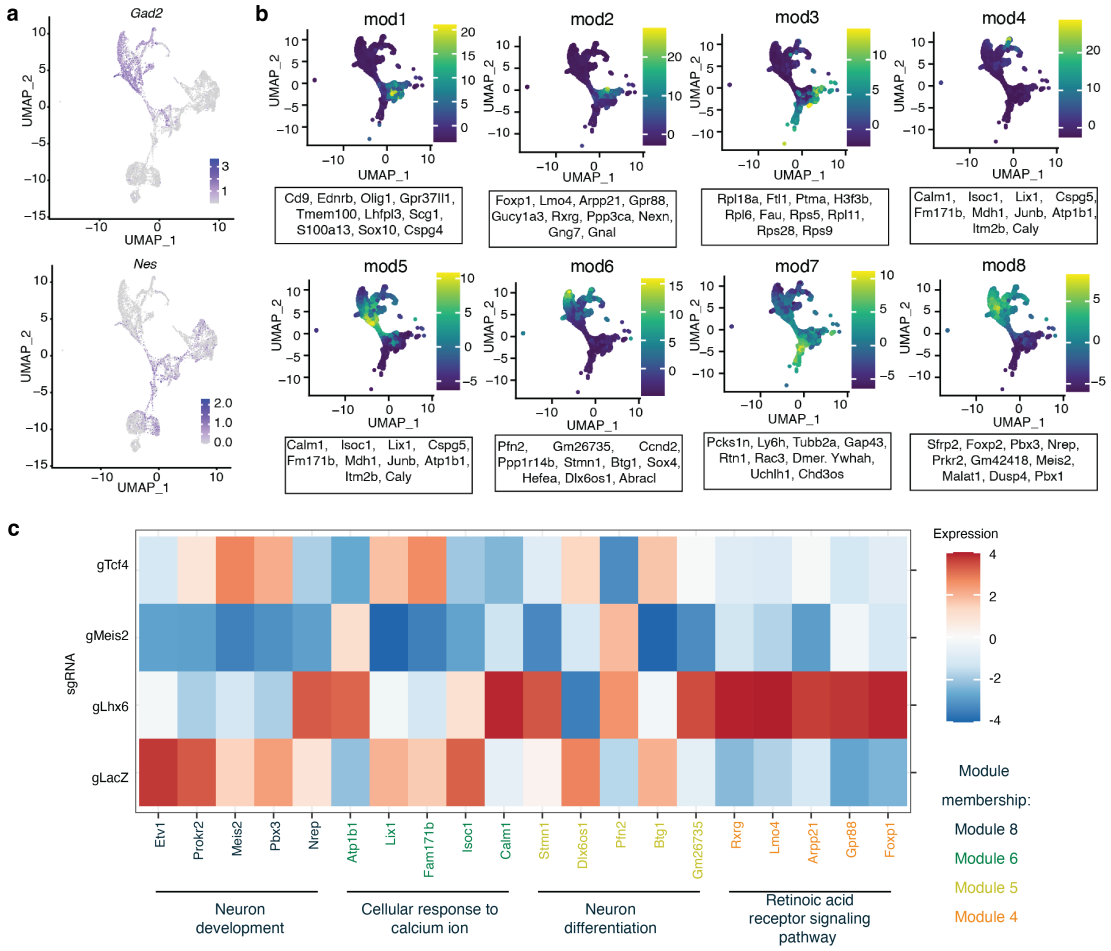


b AGACGTTGACAGTGACACCTTTCCCTCTTC~~TGT~~TTGGATAATTGCAT~~TGT~~AAGAATTA
CTTATCATACATTTAACAGTTAATACCTATTTATTATGAGTTCATGTAACTGCTATTTAA
GGAATAATCGGCACTAACTCCTCTCTCCTCCACTTGATGCCACTT~~TGA~~TTCTGTAATT
GCTGGTTTCTGCGATATTATGCCACAGAATC~~TGA~~AAGCATAATTGGTTGTTTCATTGT
AGACAATAC

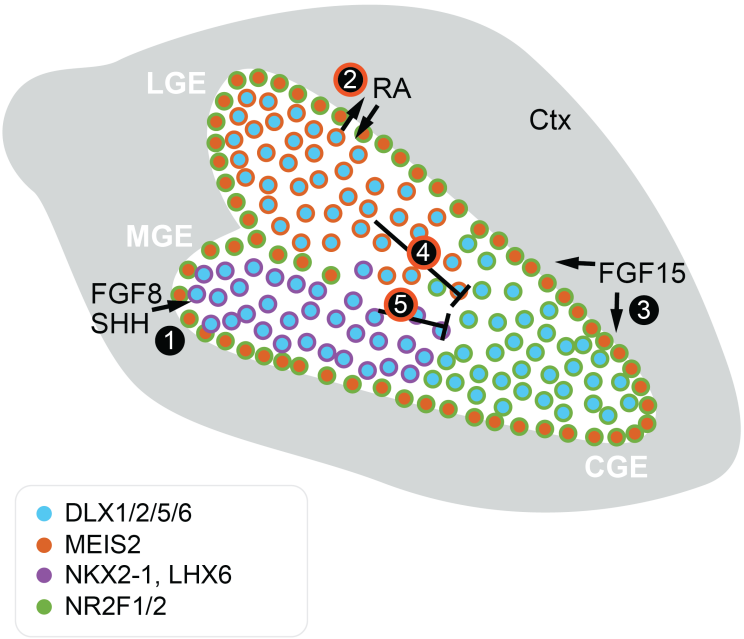
Supplementary Figure 5: MEIS2 and DLX5 cooperatively regulate the *Drd1* gene locus. a, The 4C data of the *Drd1* locus from Hi-C data of NPC is shown with aligned tracks of MEIS1/2 ChIP-seq at E14.5 (red), DLX5 ChIP-seq at E13.5 (blue), LGE scATAC-seq at E12.5 (dark grey) and MGE scATAC-seq at E12.5 (grey). **b**, The DNA sequence of the short version of the enhancer enhD1 is shown with highlighted combined MEIS (red) and DLX (blue) binding motifs. The removed TG bases in the mutated version of enhD1 are crossed out.



Supplementary Figure 6: Spatial nature of MEIS2 and its targets. **a**, In Situ Hybridization (IHC) images of *Lhx6*, *Meis2*, *Dlx5*, and *Tcf4* from the Allen Brain Institute's Developing Mouse Brain Atlas at E13.5. **b**, Overlap between binding sites of MEIS1/2, DLX5, and LHX6 with activity of Vista enhancers at E11.5 (Visel et al., 2007). **c**, Selected Vista enhancers with in vivo activity and co-binding of MEIS-DLX5-LHX6 at E11.5. MGE (medial ganglionic eminence) and LGE (lateral ganglionic eminence) are indicated.



Supplementary Figure 7: Module analysis of P7 tCROP-seq. a, Feature plots of canonical marker genes *Gad2* and *Nes* at postnatal day 7. The expression is depicted from low (gray) to high (pruple). **b**, Feature plots of gene module expression scores and the correlated genes within each module. **c**, Average expression of top 5 module genes for each sgRNA at postnatal day 7.



Supplementary Figure 8: Schematic summary of spatial factors in the ganglionic eminence leading to specific enhancer activation. MGE (medial ganglionic eminence), CGE (caudal ganglionic eminence), and LGE (lateral ganglionic eminence) are indicated. Ctx represents the cortex, RA is retinoic acid, SHH is sonic hedgehog and FGF is fibroblast growth factor. Black circles labeled (1) (Storm et al., 2006; Molotkova et al., 2007), (2) (Chatzi et al., 2011), (3) (Borello et al., 2008; Shohayeb et al., 2021), (4) (Su et al., 2022), (5) (Vogt et al., 2014) refer to previous findings. The red circle outline represents the findings of the current study.

Supplemental Material

Tables are presented as individual Excel files.

- Table S1: Selected sgRNAs list with primers plus TrackerSeq primers.
- Table S2: E16-tCrop-seq top10 marker genes.
- Table S3: E16-tCrop-seq pseudo-bulk differential gene expression analysis (DEG).
- Table S4: E16-tCrop-seq interneuron clusters differential gene expression analysis (DEG).
- Table S5: ChIP-seq supplementary information.
- Table S6: P7-tCrop-seq top10 marker genes.
- Table S7: P7-tCrop-seq bulk differential gene expression analysis.
- Table S8: Information about cloned regulatory elements, used in luciferase reporter assays.
- Table S9: Dataset summary.

Methods

Cell line

Mouse Neuro2a neuroblastoma cells (ECACC, 89121404) were cultured in Dulbecco's modified Eagle medium (DMEM, Sigma, D6429) with high glucose, L-glutamine, and sodium pyruvate supplemented with 10% (v/v) fetal bovine serum (FBS, Sigma, F9665) and containing 1% (v/v) antibiotics (100 U/mL penicillin, 100 mg/mL streptomycin) (Sigma, P0781). Neuro2a cells were incubated at 37 °Cin a 5% CO_2 humidified atmosphere and passaged twice a week. Cell passage numbers were limited to no more than 10.

sgRNA vector construction

The piggyBac based backbone plasmid contains sgRNAs under mouse U6 promoter, a gift from Randy Platt, were modified by adding pCAG-TdTomato (Addgene, 59462) and a capture sequence at the scaffold of sgRNA for 10x feature barcode retrieval (cs1 incorporated at the 3' end; (Replogle et al., 2020)) with use of NEBuilder HiFi DNA Assembly (NEB, E2621). sgRNAs were designed using CRISPick for CRISPRko (Doench et al., 2016; Sanson et al., 2018) and validated with INDELPHI (Shen et al., 2018) for high frame shift efficiency. At least 3 sgRNAs per gene were cloned using ssDNAs oligoes (IDT) and NEBuilder HiFi DNA Assembly (NEB) into modified backbone. The efficiency of sgRNA was measured in Neuro2A cells. Cells were transfected with pCAG-Cas9-EGFP (gift from Randy Platt) and sgRNAs plasmids with TransIT-LT1 Transfection Reagent (Mirus, MIR2305) and after 48 h were sorted with D FACSAria III Cell Sorter (BD FACSDiva Software, version 8.0.2) for TdTomato and EGFP. The genomic DNA was extracted with Quick-DNA Miniprep Plus Kit (Zymo, D4068) and the region around sgRNA targeting was amplified with Q5 polymerase (NEB, M094S) with primers listed in the Table S1, and afterwards send for Sanger sequencing at Microsynth Seqlab GmbH. The knockout efficiency quantified using the Synthego ICE Analysis Tool (Hsiau et al., 2019). The results for selected sgRNAs are shown in the Table S1.

TrackerSeq library preparation and validation

TrackerSeq is a piggyBac transposon-based (Ding et al., 2005) library, developed to be compatible with the 10x single-cell transcriptomic platform (Bandler et al., 2022). It records the in vivo lineage history of single cells through the integration of multiple oligonucleotide sequences into the mouse genome. Each of these individual lineage barcodes is a 37-bp long synthetic nucleotide that consists of short random nucleotides bridged by fixed nucleotides. This design results in a library with a theoretical complexity of approximately 4.3 million lineage barcodes (16^8) with each barcode differing from another by at least 5 bp. To construct the library, the piggyBac donor

plasmid (Addgene, 40973) was altered to include a number of modifications (Bandler et al., 2022). A Read2 partial primer sequence was cloned into the 3' UTR of the EGFP to enable retrieval by the 10x platform. The sucrose gene was cloned into the vector, so that empty plasmids that fail to incorporate a lineage barcode during the cloning process are removed. Following digestion with BstXI (Jena Bioscience, EN-E2118) to remove the sucrose gene, the plasmid was run on a gel and column purified. The lineage barcode oligo mix was cloned downstream of the Read2 partial primer sequence in the purified donor plasmid via multiple Gibson Assembly reactions (NEB, E2611S). Gibson assembly reactions were then pooled and desalted with 0.025 μm MCE membrane (Millipore, VSWP02500) for 40 min, and finally concentrated using a SpeedVac. 3 μl of the purified assembly is incubated with 50 μl of NEB 10-β-competent E.coli cells (NEB, C3019H) for 30 min at 4 °C, then electroporated at 2.0 kV, 200 Ω, 25 μF (Bio-Rad, Gene Pulser Xcell Electroporation Systems). Electroporated E.coli were incubated for 90 min shaking at 37 °Cand then plated into pre-warmed sucrose/ampicillin plates. The colonies were scraped off the plates after 8 h, and the plasmids were grown in LB medium with ampicillin up to OD = 0.5. The plasmid library was purified using a column purification kit (Zymo, D4202). We first assessed the integrity of the TrackerSeq barcode libraries by sequencing the library to a depth of approximately 42 million reads to test whether any barcode was over-represented. Around 3.6 million valid lineage barcodes that had a quality score of 30 or higher were extracted from the R2 FASTQ files using Bartender (Zhao et al., 2018). One thousand barcodes were randomly sampled from the extracted lineage barcodes to assess hamming distance. To group similar extracted barcodes into putative barcodes, Bartender assigns a UMI to each barcode read to handle PCR jackpotting errors, and clusters them. The cluster distance was set to 3 so that extracted barcodes within 3 bp of each other have a chance of being clustered together. A total of 2 × 105 clusters of barcodes were identified, suggesting that the barcode library has a diversity that is at least in the 105 range.

Mice and in utero surgeries

All mouse colonies were maintained in accordance with protocols approved by the Bavarian government at the Max Planck Institute for Biological Intelligence. C57BL/6 *wt* females were crossed to C57BL/6 *wt* or to CAS9-EGFP (B6.Gt(ROSA)26Sortm1.1(CAG-cas9*,-EGFP)Fezh/J, Jax 026179) males (Platt et al., 2014). Embryos were staged in days post-coitus, with E0.5 defined as 12:00 of a day than a vaginal plug was detected after overnight mating. Timed pregnant mice were anesthetized with isoflurane (5% induction, 2.5% during the surgery) and treated with the analgesic Metamizol (WDT). A microsyringe pump (Nanoject III Programmable Nano-liter Injector, 100/240V, DRUM3-000-207) was used to inject approximately 700 nl of DNA plasmid solution made of 0.6 μg/μl pEF1a-pBase (piggyBac-transposase; a gift from R. Platt) and the sgRNA plasmid 0.7 μg/μl, diluted in endo-free TE buffer and 0.002% Fast Green FCF (Sigma, F7252), into the lateral ventricle. pCAG-Cas9-EGFP (a gift from Randy Platt) plasmid was added when *wt* males were used for plugs. Embryos were then electroporated by holding each head between platinum-plated tweezer electrodes (5 mm in diameter, BTX, 45-0489) across the uterine wall, while 5 electric pulses (35 V, 50 ms at 1 Hz) were delivered with a square-wave electroporator (BTX, ECM830) (Saito, 2006). Pregnant dams were kept in single cages and pups were kept with their mothers.

Sample collection

We collected electroporated brains from mouse embryos at E16.5 in ice-cold Leibovitz's L-15 Medium (ThermoFisher, 11415064) with 5% FBS or at P7-8 in ice-cold Hibernate-A Medium (ThermoFisher, A1247501) with 10% FBS and B-27 Supplement (ThermoFisher, 17504044). The same mediums were used during flow cytometry sorting. Only forebrains were collected, thus excluding the thalamus, hypothalamus, brainstem and cerebellum. Papain dissociation system (Wortington, LK003150) was carried out according to the protocol described in Jin et al., 2020 (Jin et al., 2020) on the gentleMACS™ Octo Dissociator (Miltenyi Biotec). To isolate positive cells, flow cytometry was done using a BD FACSAria III Cell Sorter (BD FACSDiva Software,

502 version 8.0.2) with a 100-µm nozzle. EGFP and TdTomato-positive cells were collected in bulk
503 for downstream processing on the 10x Genomics Chromium platform. After sorting 5,000–16,000
504 individual cells per sample, in PBS (Lonza) with 0.02% BSA (NEB), were loaded onto a 10X
505 Genomics Chromium platform for Gel Beads-in-emulsion (GEM) and cDNA generation carrying
506 cell- and transcript-specific barcode using the Chromium Single Cell 3' Reagent Kit v3.1 with
507 Feature Barcoding technology (PN-1000121) following manufacture protocol (document number
508 CG000205, 10X Genomics).

509 Preparation of tCROP-seq libraries

510 Uniquely barcoded RNA transcripts (poly(A)-mRNA and sgRNA) were reverse transcribed. 3'
511 Gene Expression library and CRISPR Screening library were generated according to manufacturer's
512 user guide (Document number CG000205) with use of Chromium Library v3.1 kit (PN-1000121),
513 Feature Barcode Library Kit (PN-1000079) and Single Index Kit (PN-1000213) (10X Genomics).
514 Libraries were quantified with Agilent BioAnalyzer.

515 Preparation of TrackerSeq NGS libraries

516 The TrackerSeq lineage libraries retrieved from RNA were amplified with Q5 polymerase (NEB,
517 M094S) in a 50-µl reaction, using 10 µl of cDNA as template (Bandler et al., 2022). Specifically,
518 each PCR contained: 25 µl Q5 High-fidelity 2X Master Mix, 2.5 µl 10 µM P7-indexed reverse
519 primer, 2.5 µl 10 µM i5-indexed forward primer, 10 µl molecular grade H_2O, 10 µl cDNA (for
520 primer sequences and indices, see Table S1). Libraries were purified with a dual-sided SPRI
521 selection using Beckman Coulter Agencourt RNAClean XP (A63987), and quantified with an
522 Agilent BioAnalyzer.

Sequencing and read mapping

Transcriptome and CRISPR barcode libraries were sequenced either on an Illumina NextSeq 500 at the Next Generation Sequencing Facility of the Max Planck Institute of Biochemistry or on a NovaSeq at the Genomics Core Facility at the Helmholtz Center in Munich. For a detailed report on each dataset, see Table S9. Sequencing reads in FASTQ files were aligned to a reference transcriptome (mm10-2.1.0) and collapsed into UMI counts using the 10x Genomics Cell Ranger software (version 3.0.2 or 5.0.1).

Cell filtering, data normalization, batch correction and clustering of tCROP-seq datasets

Processing embryonic tCROP-seq datasets. Embryonic E16 tCROP-seq datasets, including those that contained TrackerSeq barcodes, were processed together for cell filtering, data normalization and cluster annotation following the standard Seurat workflow (4.0.6, (Hafemeister and Satija, 2019)). Data was read into R as a count matrix. Each dataset was preprocessed according to a set of criterions: minimum and maximum genes expressed, maximum nCount_RNA, and mitochondrial mapping percentage. CRISPR-perturbed cells were identified using a CSV file output by CellRanger that contained the cell barcodes and the sgRNA detected in that cell. We removed excitatory clusters by removing those that have UMI counts for Neurod2>2 and Neurod6>2, which are markers of excitatory neurons. To create an 'integrated' data assay, we combined the embryonic tCROP-seq dataset with *wt* GE datasets that were collected at E13.5 and E15.5 as described by Stuart et al. (Stuart et al., 2019). Briefly, after each dataset is normalized using SCTransoform(), anchors are identified using FindIntegrationAnchors() then the anchors are used to integrate the embryonic tCROP-seq and *wt* datasets with IntegrateData().

To group cells into clusters, we first constructed a shared-nearest neighbour using the FindNeighbors() algorithm, then input the graph into an SLM algorithm that is implemented through the FindClusters() function in Seurat (dimensions = 30, res = 0.8). We obtained cluster-specific marker

genes by performing differential expression analysis using FindAllMarkers(), comparing cells of each cluster to cells from all other clusters. Genes were considered differentially expressed if they met the fold change, minimum expression and adjusted P value cut-offs as dictated by the Wilcoxon rank sum test implemented via Seurat. Clusters were assigned to cell types based on marker gene expression from literature, primarily http://mousebrain.org/development/ (La Manno et al., 2021).

Processing postnatal tCROP-seq datasets. Each post-natal P7 dataset was preprocessed according to a set of criterion: minimum and maximum genes expressed, maximum nCount_RNA, and mitochondrial mapping percentage. We normalized and regressed out technical effects introduced by nFeature_RNA, nCount_RNA, and mitochondrial mapping percentage using SCTransform().

We used Harmony (v1.0, (Korsunsky et al., 2019)) within the Seurat workflow using default settings (theta = 2, lambda = 1, sigma = 0.1) to integrate different CRISPR datasets. We used the first 30 Harmony embeddings for UMAP (https://github.com/lmcinnes/umap) visualizations and clustering analysis.

To group cells into clusters, we first constructed a shared-nearest neighbour graph from Harmony embeddings using the FindNeighbors() algorithm, then input the graph into the FindClusters() function in Seurat (dimensions = 30, res = 0.8). We obtained cluster-specific marker genes by performing differential expression analysis using FindAllMarkers().

Comparing cell type composition between perturbations

We compared the perturbation effect on cell type composition using the method described by Jin et al.(Jin et al., 2020). More specifically, we used a poisson regression model to test the relationship between perturbations and cell number, correcting for batch and total number of cells. The formula is as follows:

Num offset(logTot) + Batch + Pert

Differential expression analysis

We used Libra package to perform DEA (Squair et al., 2021). We ran the run_DE functions on Seurat objects directly with the following parameters (de_family = pseudobulk, de_family = pseudobulk, de_method = edgeR, de_type = LRT). We obtained DEGs of PNs or INs by using run_DE function on cells grouped into classes (mitotic, projection neurons, and interneurons). To obtain DEGs of individual subclusters, we used the run_DE function on individual clusters. We filtered for statistically significant genes (FDR-adjusted p-value threshold = 0.05). Genes were considered differentially expressed if avg_logFC < -1.0 or avg_logFC > 1.0.

Hotspot analysis of gene coexpression

Hotspot(v0.91) is a tool for identifying co-expressing gene modules in a single-cell dataset (DeTomaso and Yosef, 2021). It computes gene modules by evaluating the pairwise correlation of genes with high local autocorrelation, then clusters the results into a gene-gene affinity matrix. *Gad2*-expressing inhibitory population in the P7 dataset was first subset out from the rest to identify inhibitory specific modules in the embryonic dataset. We ran the depth-adjusted negative binomial model on the entire count matrix and Harmony (v1.0) corrected principal components (see batch correction of tCROP-seq datasets). We computed a k-nearest-neighbors (KNN) graph with 30 neighbours, 9154 non-varying genes were subsequently detected and removed. Autocorrelations between each gene were calculated, and the top 500 significant (FDR < 0.05) genes were used to evaluate pairwise gene associations (local correlations). After pairwise local correlations are calculated, we grouped genes into modules. Modules were created through agglomerative clustering, where the minimum number of genes per module was set to 30. 8 modules were identified, and 103 genes were not assigned to a module. Summary per-cell module scores are calculated using the calculate_module_scores() function as described by DeTomaso et al. (DeTomaso and Yosef, 2021).

Testing Hotpot module gene sets

As described by Jin et al.(Jin et al., 2020), linear regression was used to test the relationship between perturbation and Hotspot module gene scores. The batch and number of genes were corrected for using the lm function in R, with the following formula:

Gene Score perturbation + batch + nGene

GO Term analysis of differentially expressed genes and module genes

GO Term analysis was done using the package enrichR (Kuleshov et al., 2016). The DEGs and module genes of each module were queried against the following databases: GO_Molecular_Functio_2018, GO_Cellular_Component_2018, and GO_Biological_Process_2018. Only GO Terms that were significant (p-value adjusted < 0.05) were kept.

TrackerSeq barcode processing and analysis

We followed the protocol outlined in (Bandler et al., 2022) to process the TrackerSeq barcodes and obtain cloneIDs for each corresponding cell barcode. The resulting cloneIDs were added to the Seurat object metadata of these tCROP-seq datasets: ED210204, ED210215, ED211111, ED211124 for lineage analysis. To quantify clonal relationship between cell classes, the inhibitory clusters were first merged into cell classes (Figure 2c) based on whether they were annotated as mitotic (*Ube2c* and *Top2a*), or as INs and PNs (*Gad2*). The UpsetR library was used to count the number of clones shared between the neuronal classes, as well as the proportion of clonal relationships in gMeis2 and gLacZ datasets. The set size is the number of cells in the class. The UpSet bar plot show the calculated proportion of each type of clonal distribution category within the perturbation. Each percentage was obtained by dividing the clones belonging to that category (e.g. clones containing only mitotic and INs) by the number of clones belonging to all other categories of clonal distribution.

Luciferase assay

Regulatory elements were amplified from mouse genomic DNA with Q5 polymerase (NEB, M0491) using primers listed in Table S8 and cloned into pGL4.24[luc2P/minP] (Promega, E8421) with NEBuilder HiFi DNA Assembly kit (NEB, E2621). Enhancer hs1080 had to be cloned in reverse-complement. Mouse *Meis2* isoform D (4) (the tag was removed) and *Lhx6* variant 1 (C-DYK) expressing vectors were purchased from Genscript, *Dlx5* and *Pbx1* genes were amplified from mouse cDNA and cloned into pcDNA3.1 (Genscript). Meis2 vector was mutated with NEBuilder HiFi DNA Assembly kit (NEB, E2621) to harbor the human mutation p.(Arg333Lys), c.998G>A (Meis2*333) (Verheije et al., 2019). A short version of enhD1 luciferase vector was mutated with use of a gBlock (IDT) and NEBuilder HiFi DNA Assembly kit (NEB, E2621). Luciferase reporter vectors were co-transfected with pNL1.1.PGK[Nluc/PGK] (Promega), pcDNA3 or pcDNA3-Dlx5, pcDNA3-PBX1, pcDNA3-Meis2, pcDNA3-Lhx6. Neuro2a cells were seeded in 24-well plates at 80,000 cells per well and on the next day were transfected with TransIT-LT1 Transfection Reagent (Mirus, MIR 2300), using 150 ng luciferase reporter, 10 ng Nluc/PGK and 350 ng total of pcDNA3 plasmids per well (150 ng per TFs vector). Cells were harvested 24 hours after transfection and luciferases activity was measured using Nano-Glo® Dual-Luciferase® Reporter Assay System (Promega) on Berthold Multimode reader Tristar2S. A Nanoluc reporter was used for normalization.

Chromatin-Immunoprecipitation (ChIP)

Mice were handled in accordance with the CNIC Ethics Committee, Spanish laws, and the EU Directive 2010/63/EU for the use of animals in research. GEs and part of the underlying striatum of 70 *wt* C57BL/6 embryos at E14.5 were microdissected and immediately fixed in 1% formaldehyde for 5 min. Tissue preparation, immunoprecipitation and sequencing on an Illumina HiSeq2500 were performed as previously described (Delgado et al., 2021). Immunoprecipitation was carried out using a combination of two anti-MEIS antibodies, one recognizing MEIS1A and MEIS2A, the other recognizing all MEIS2 isoforms (Mercader et al., 2005).

ChIP-seq Data Analysis

61 bp single-end reads were trimmed using Cutadapt (v1.16) and mapped to GRCm38 using Bowtie2 (v2.3.0) (Langmead and Salzberg, 2012) followed by peak calling with MACS2 (v2.1.2) (Feng et al., 2012) using a cutoff of q=0.01. TSS definitions were adapted from the eukaryotic promoter database (mmEPDnew version 003) (Meylan et al., 2020). We determined the distance of each peak to the nearest TSS using the R package plyranges (v1.180). Using custom R scripts, peaks were assigned to the TSS of a gene when overlapping a ca. 5 kb region around a TSS, defined as promoter region. Overlap with developmental enhancers (Gorkin et al., 2020) was determined in the same way. Similarly, we determined overlap of MEIS2 binding sites with DLX5 binding sites at E13.5 from Lindtner et al. (Lindtner et al., 2019) and LHX6 binding sites at E13.5 from Sandberg et al. (Sandberg et al., 2016). Enrichment of enhancer-overlapping peaks among shared MEIS2/DLX5 peaks, compared to MEIS2- and DLX5-exclusive peaks, was determined using Peason's Chi-squared test of the R 'stats' package(v4.0.2). Genomic tracks and Vista enhancers (Visel et al., 2007) were visualized using the Integrated Genomics Viewer (v2.4.1) (Robinson et al., 2011).

Motif identification and enrichment of known motifs were carried out by HOMER (v4.10.4) (Heinz et al., 2010) using default settings. Motif enrichment within enhancer- and promoter-overlapping peaks was likewise performed with HOMER. We used SpaMo (v5.4.1) (Whitington et al., 2011)) to determine motif spacing between MEIS2 and DLX5 binding motifs in common MEIS2/DLX5 binding sites, within 100 bp up- and down- stream of MEIS2 peak summits.

Data used in this study

GSE167047 (snATAC-seq of E12.5 MGE and LGE; (Rhodes et al., 2022)), GSE85705 (LHX6-ChIP-seq GE E13.5; (Sandberg et al., 2016)), GSE124936 (DLX5-ChIP-seq GE E13.5; (Lindtner et al., 2019)) and GSE188528 (scRNA-seq of LGE, MGE, CGE E13.5; (Bandler et al., 2022)) were downloaded from GEO (https://www.ncbi.nlm.nih.gov/geo/). Developmental enhancers and interacting genes (Gorkin et al., 2020). Vista enhancers are from Vista enhancer

browser (https://enhancer.lbl.gov) (Visel et al., 2007). ChIP-seq and ATAC-seq tracks were presented using IGV software, which is publicly available from the Broad Institute website (https://software.broadinstitute.org/software/igv/). Hi-C data from NPC cells (Bonev et al., 2017) was visualized with 3D Genome Browser (http://3dgenome.fsm.northwestern.edu/index.html).

References

Agoston, Z., Heine, P., Brill, M. S., Grebbin, B. M., Hau, A.-C., Kallenborn-Gerhardt, W., Schramm, J., Götz, M., and Schulte, D. (2014). Meis2 is a pax6 co-factor in neurogenesis and dopaminergic periglomerular fate specification in the adult olfactory bulb. *Development*, 141(1):28–38.

Anderson, S. A., Eisenstat, D. D., Shi, L., and Rubenstein, J. L. (1997). Interneuron migration from basal forebrain to neocortex: dependence on dlx genes. *Science*, 278(5337):474–6.

Anderson, S. A., Marín, O., Horn, C., Jennings, K., and Rubenstein, J. L. (2001). Distinct cortical migrations from the medial and lateral ganglionic eminences. *Development*, 128(3):353–63.

Asgarian, Z., Oliveira, M. G., Stryjewska, A., Maragkos, I., Rubin, A. N., Magno, L., Pachnis, V., Ghorbani, M., Hiebert, S. W., Denaxa, M., and Kessaris, N. (2022). Mtg8 interacts with lhx6 to specify cortical interneuron subtype identity. *Nat Commun*, 13(1):5217.

Bandler, R. C., Mayer, C., and Fishell, G. (2017). Cortical interneuron specification: the juncture of genes, time and geometry. *Curr Opin Neurobiol*, 42:17–24.

Bandler, R. C., Vitali, I., Delgado, R. N., Ho, M. C., Dvoretskova, E., Ibarra Molinas, J. S., Frazel, P. W., Mohammadkhani, M., Machold, R., Maedler, S., Liddelow, S. A., Nowakowski, T. J., Fishell, G., and Mayer, C. (2022). Single-cell delineation of lineage and genetic identity in the mouse brain. *Nature*, 601(7893):404–409.

Batista-Brito, R., Machold, R., Klein, C., and Fishell, G. (2008). Gene expression in cortical interneuron precursors is prescient of their mature function. *Cereb Cortex*, 18(10):2306–17.

Berenguer, M. and Duester, G. (2022). Retinoic acid, rars and early development. *J Mol Endocrinol*, 69(4):T59–T67.

Bonev, B., Mendelson Cohen, N., Szabo, Q., Fritsch, L., Papadopoulos, G. L., Lubling, Y., Xu, X., Lv, X., Hugnot, J.-P., Tanay, A., and Cavalli, G. (2017). Multiscale 3d genome rewiring during mouse neural development. *Cell*, 171(3):557–572.e24.

Borello, U., Cobos, I., Long, J. E., McWhirter, J. R., Murre, C., and Rubenstein, J. L. R. (2008). Fgf15 promotes neurogenesis and opposes fgf8 function during neocortical development. *Neural Dev*, 3:17.

Bridoux, L., Zarrineh, P., Mallen, J., Phuycharoen, M., Latorre, V., Ladam, F., Losa, M., Baker, S. M., Sagerstrom, C., Mace, K. A., Rattray, M., and Bobola, N. (2020). Hox paralogs selectively convert binding of ubiquitous transcription factors into tissue-specific patterns of enhancer activation. *PLoS Genet*, 16(12):e1009162.

Cesario, J. M., Landin Malt, A., Deacon, L. J., Sandberg, M., Vogt, D., Tang, Z., Zhao, Y., Brown, S., Rubenstein, J. L., and Jeong, J. (2015). Lhx6 and lhx8 promote palate development through negative regulation of a cell cycle inhibitor gene, p57kip2. *Hum Mol Genet*, 24(17):5024–39.

Chang, C. P., Jacobs, Y., Nakamura, T., Jenkins, N. A., Copeland, N. G., and Cleary, M. L. (1997). Meis proteins are major in vivo dna binding partners for wild-type but not chimeric pbx proteins. *Mol Cell Biol*, 17(10):5679–87.

Chatzi, C., Brade, T., and Duester, G. (2011). Retinoic acid functions as a key gabaergic differentiation signal in the basal ganglia. *PLoS Biol*, 9(4):e1000609.

Datlinger, P., Rendeiro, A. F., Schmidl, C., Krausgruber, T., Traxler, P., Klughammer, J., Schuster, L. C., Kuchler, A., Alpar, D., and Bock, C. (2017). Pooled crispr screening with single-cell transcriptome readout. *Nat Methods*, 14(3):297–301.

Delgado, I., Giovinazzo, G., Temiño, S., Gauthier, Y., Balsalobre, A., Drouin, J., and Torres, M. (2021). Control of mouse limb initiation and antero-posterior patterning by meis transcription factors. *Nat Commun*, 12(1):3086.

den Hoed, J., Devaraju, K., and Fisher, S. E. (2021). Molecular networks of the foxp2 transcription factor in the brain. *EMBO Rep*, 22(8):e52803.

DeTomaso, D. and Yosef, N. (2021). Hotspot identifies informative gene modules across modalities of single-cell genomics. *Cell Syst*, 12(5):446–456.e9.

Ding, S., Wu, X., Li, G., Han, M., Zhuang, Y., and Xu, T. (2005). Efficient transposition of the piggybac (pb) transposon in mammalian cells and mice. *Cell*, 122(3):473–83.

Doench, J. G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E. W., Donovan, K. F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., Virgin, H. W., Listgarten, J., and Root, D. E. (2016). Optimized sgrna design to maximize activity and minimize off-target effects of crispr-cas9. *Nat Biotechnol*, 34(2):184–191.

Douglas, G., Cho, M. T., Telegrafi, A., Winter, S., Carmichael, J., Zackai, E. H., Deardorff, M. A., Harr, M., Williams, L., Psychogios, A., Erwin, A. L., Grebe, T., Retterer, K., and Juusola, J. (2018). De novo missense variants in meis2 recapitulate the microdeletion phenotype of cardiac and palate abnormalities, developmental delay, intellectual disability and dysmorphic features. *Am J Med Genet A*, 176(9):1845–1851.

Dupacova, N., Antosova, B., Paces, J., and Kozmik, Z. (2021). Meis homeobox genes control progenitor competence in the retina. *Proc Natl Acad Sci U S A*, 118(12).

Feng, J., Liu, T., Qin, B., Zhang, Y., and Liu, X. S. (2012). Identifying chip-seq enrichment using macs. *Nat Protoc*, 7(9):1728–40.

Flames, N., Pla, R., Gelman, D., Rubenstein, J., Puelles, L., and Marín, O. (2007). Delineation of multiple subpallial progenitor domains by the combinatorial expression of transcriptional codes. *J. Neurosci.; Journal of Neuroscience*, 27(36):9682–9695.

French, C. A. and Fisher, S. E. (2014). What can mice tell us about foxp2 function? *Curr Opin Neurobiol*, 28:72–9.

Gangfuß, A., Yigit, G., Altmüller, J., Nürnberg, P., Czeschik, J. C., Wollnik, B., Bögershausen, N., Burfeind, P., Wieczorek, D., Kaiser, F., Roos, A., Kölbel, H., Schara-Schmidt, U., and Kuechler, A. (2021). Intellectual disability associated with craniofacial dysmorphism, cleft palate, and congenital heart defect due to a de novo meis2 mutation: A clinical longitudinal study. *Am J Med Genet A*, 185(4):1216–1221.

Gelman, D., Griveau, A., Dehorter, N., Teissier, A., Varela, C., Pla, R., Pierani, A., and Marín, O. (2011). A wide diversity of cortical gabaergic interneurons derives from the embryonic preoptic area. *J Neurosci*, 31(46):16570–80.

Gerfen, C. R. and Surmeier, D. J. (2011). Modulation of striatal projection systems by dopamine. *Annu Rev Neurosci*, 34:441–66.

Giliberti, A., Currò, A., Papa, F. T., Frullanti, E., Ariani, F., Coriolani, G., Grosso, S., Renieri, A., and Mari, F. (2020). Meis2 gene is responsible for intellectual disability, cardiac defects and a distinct facial phenotype. *Eur J Med Genet*, 63(1):103627.

Gorkin, D. U., Barozzi, I., Zhao, Y., Zhang, Y., Huang, H., Lee, A. Y., Li, B., Chiou, J., Wildberg, A., Ding, B., Zhang, B., Wang, M., Strattan, J. S., Davidson, J. M., Qiu, Y., Afzal, V., Akiyama, J. A., Plajzer-Frick, I., Novak, C. S., Kato, M., Garvin, T. H., Pham, Q. T., Harrington, A. N., Mannion, B. J., Lee, E. A., Fukuda-Yuzawa, Y., He, Y., Preissl, S., Chee, S., Han, J. Y., Williams, B. A., Trout, D., Amrhein, H., Yang, H., Cherry, J. M., Wang, W., Gaulton, K., Ecker, J. R., Shen, Y., Dickel, D. E., Visel, A., Pennacchio, L. A., and Ren, B. (2020). An atlas of dynamic chromatin landscapes in mouse fetal development. *Nature*, 583(7818):744–751.

Hafemeister, C. and Satija, R. (2019). Normalization and variance stabilization of single-cell rna-seq data using regularized negative binomial regression. *Genome Biol*, 20(1):296.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H., and Glass, C. K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and b cell identities. *Mol Cell*, 38(4):576–89.

Hobert, O. and Westphal, H. (2000). Functions of lim-homeobox genes. *Trends Genet*, 16(2):75–83.

Hsiau, T., Conant, D., Rossi, N., Maures, T., Waite, K., Yang, J., Joshi, S., Kelso, R., Holden, K., Enzmann, B., and Stoner, R. (2019). Inference of crispr edits from sanger trace data. *bioRxiv*.

Hyman-Walsh, C., Bjerke, G. A., and Wotton, D. (2010). An autoinhibitory effect of the homothorax domain of meis2. *FEBS J*, 277(12):2584–97.

Jin, X., Simmons, S. K., Guo, A., Shetty, A. S., Ko, M., Nguyen, L., Jokhi, V., Robinson, E., Oyler, P., Curry, N., Deangeli, G., Lodato, S., Levin, J. Z., Regev, A., Zhang, F., and Arlotta, P. (2020). In vivo perturb-seq reveals neuronal and glial abnormalities associated with autism risk genes. *Science*, 370(6520).

Jindal, G. A. and Farley, E. K. (2021). Enhancer grammar in development, evolution, and disease: dependencies and interplay. *Dev Cell*, 56(5):575–587.

Jolma, A., Yin, Y., Nitta, K. R., Dave, K., Popov, A., Taipale, M., Enge, M., Kivioja, T., Morgunova, E., and Taipale, J. (2015). Dna-dependent formation of transcription factor pairs alters their binding specificity. *Nature*, 527(7578):384–8.

Kim, H., Berens, N. C., Ochandarena, N. E., and Philpot, B. D. (2020). Region and cell type distribution of tcf4 in the postnatal mouse brain. *Front Neuroanat*, 14:42.

Knowles, R., Dehorter, N., and Ellender, T. (2021). From progenitors to progeny: Shaping striatal circuit development and function. *J Neurosci*, 41(46):9483–9502.

Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.-R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with harmony. *Nat Methods*, 16(12):1289–1296.

Kreitzer, A. C. and Malenka, R. C. (2008). Striatal plasticity and basal ganglia circuit function. *Neuron*, 60(4):543–54.

Kuerbitz, J., Arnett, M., Ehrman, S., Williams, M. T., Vorhees, C. V., Fisher, S. E., Garratt, A. N., Muglia, L. J., Waclaw, R. R., and Campbell, K. (2018). Loss of intercalated cells (itcs) in the mouse amygdala of tshz1 mutants correlates with fear, depression, and social interaction phenotypes. *J Neurosci*, 38(5):1160–1177.

Kuleshov, M. V., Jones, M. R., Rouillard, A. D., Fernandez, N. F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S. L., Jagodnik, K. M., Lachmann, A., McDermott, M. G., Monteiro, C. D., Gundersen, G. W., and Ma'ayan, A. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*, 44(W1):W90–7.

La Manno, G., Siletti, K., Furlan, A., Gyllborg, D., Vinsland, E., Mossi Albiach, A., Mattsson Langseth, C., Khven, I., Lederer, A. R., Dratva, L. M., Johnsson, A., Nilsson, M., Lönnerberg, P., and Linnarsson, S. (2021). Molecular architecture of the developing mouse brain. *Nature*, 596(7870):92–96.

Langmead, B. and Salzberg, S. L. (2012). Fast gapped-read alignment with bowtie 2. *Nat Methods*, 9(4):357–9.

Leung, R. F., George, A. M., Roussel, E. M., Faux, M. C., Wigle, J. T., and Eisenstat, D. D. (2022). Genetic regulation of vertebrate forebrain development by homeobox genes. *Front Neurosci*, 16:843794.

Lim, L., Mi, D., Llorca, A., and Marín, O. (2018). Development and functional diversification of cortical interneurons. *Neuron*, 100(2):294–313.

Lindtner, S., Catta-Preta, R., Tian, H., Su-Feher, L., Price, J. D., Dickel, D. E., Greiner, V., Silberberg, S. N., McKinsey, G. L., McManus, M. T., Pennacchio, L. A., Visel, A., Nord, A. S., and Rubenstein, J. L. R. (2019). Genomic resolution of dlx-orchestrated transcriptional circuits driving development of forebrain gabaergic neurons. *Cell Rep*, 28(8):2048–2063.e8.

Longobardi, E., Penkov, D., Mateos, D., De Florian, G., Torres, M., and Blasi, F. (2014). Biochemistry of the tale transcription factors prep, meis, and pbx in vertebrates. *Dev Dyn*, 243(1):59–75.

Louw, J. J., Corveleyn, A., Jia, Y., Hens, G., Gewillig, M., and Devriendt, K. (2015). Meis2 involvement in cardiac development, cleft palate, and intellectual disability. *Am J Med Genet A*, 167A(5):1142–6.

Mayer, C., Hafemeister, C., Bandler, R. C., Machold, R., Batista Brito, R., Jaglin, X., Allaway, K., Butler, A., Fishell, G., and Satija, R. (2018). Developmental diversification of cortical inhibitory interneurons. *Nature*, 555(7697):457–462.

Mercader, N., Tanaka, E. M., and Torres, M. (2005). Proximodistal identity during vertebrate limb regeneration is regulated by meis homeodomain proteins. *Development*, 132(18):4131–42.

Mesman, S., Bakker, R., and Smidt, M. P. (2020). Tcf4 is required for correct brain development during embryogenesis. *Molecular and Cellular Neuroscience*, 106:103502.

Meylan, P., Dreos, R., Ambrosini, G., Groux, R., and Bucher, P. (2020). Epd in 2020: enhanced data visualization and extension to ncrna promoters. *Nucleic Acids Res*, 48(D1):D65–D69.

Miyoshi, G., Young, A., Petros, T., Karayannis, T., McKenzie Chang, M., Lavado, A., Iwano, T., Nakajima, M., Taniguchi, H., Huang, Z. J., Heintz, N., Oliver, G., Matsuzaki, F., Machold,

R. P., and Fishell, G. (2015). Prox1 regulates the subtype-specific development of caudal ganglionic eminence-derived gabaergic cortical interneurons. *J Neurosci*, 35(37):12869–89.

Molotkova, N., Molotkov, A., and Duester, G. (2007). Role of retinoic acid during forebrain development begins late when raldh3 generates retinoic acid in the ventral subventricular zone. *Dev Biol*, 303(2):601–10.

Ng, F. S. L., Schütte, J., Ruau, D., Diamanti, E., Hannah, R., Kinston, S. J., and Göttgens, B. (2014). Constrained transcription factor spacing is prevalent and important for transcriptional control of mouse blood cells. *Nucleic Acids Res*, 42(22):13513–24.

Panganiban, G. and Rubenstein, J. L. R. (2002). Developmental functions of the distal-less/dlx homeobox genes. *Development*, 129(19):4371–86.

Platt, R. J., Chen, S., Zhou, Y., Yim, M. J., Swiech, L., Kempton, H. R., Dahlman, J. E., Parnas, O., Eisenhaure, T. M., Jovanovic, M., Graham, D. B., Jhunjhunwala, S., Heidenreich, M., Xavier, R. J., Langer, R., Anderson, D. G., Hacohen, N., Regev, A., Feng, G., Sharp, P. A., and Zhang, F. (2014). Crispr-cas9 knockin mice for genome editing and cancer modeling. *Cell*, 159(2):440–55.

Replogle, J. M., Norman, T. M., Xu, A., Hussmann, J. A., Chen, J., Cogan, J. Z., Meer, E. J., Terry, J. M., Riordan, D. P., Srinivas, N., Fiddes, I. T., Arthur, J. G., Alvarado, L. J., Pfeiffer, K. A., Mikkelsen, T. S., Weissman, J. S., and Adamson, B. (2020). Combinatorial single-cell crispr screens by direct guide rna capture and targeted sequencing. *Nat Biotechnol*, 38(8):954–961.

Rhodes, C. T., Thompson, J. J., Mitra, A., Asokumar, D., Lee, D. R., Lee, D. J., Zhang, Y., Jason, E., Dale, R. K., Rocha, P. P., and Petros, T. J. (2022). An epigenome atlas of neural progenitors within the embryonic mouse forebrain. *Nat Commun*, 13(1):4196.

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., and Mesirov, J. P. (2011). Integrative genomics viewer. *Nat Biotechnol*, 29(1):24–6.

Saito, T. (2006). In vivo electroporation in the embryonic mouse central nervous system. *Nat Protoc*, 1(3):1552–8.

Sandberg, M., Flandin, P., Silberberg, S., Su-Feher, L., Price, J. D., Hu, J. S., Kim, C., Visel, A., Nord, A. S., and Rubenstein, J. L. R. (2016). Transcriptional networks controlled by nkx2-1 in the development of forebrain gabaergic neurons. *Neuron*, 91(6):1260–1275.

Sandberg, M., Taher, L., Hu, J., Black, B. L., Nord, A. S., and Rubenstein, J. L. R. (2018). Genomic analysis of transcriptional networks directing progression of cell states during mge development. *Neural Dev*, 13(1):21.

Sanson, K. R., Hanna, R. E., Hegde, M., Donovan, K. F., Strand, C., Sullender, M. E., Vaimberg, E. W., Goodale, A., Root, D. E., Piccioni, F., and Doench, J. G. (2018). Optimized libraries for crispr-cas9 genetic screens with multiple modalities. *Nat Commun*, 9(1):5416.

Schulte, D. and Geerts, D. (2019). Meis transcription factors in development and disease. *Development*, 146(16).

Selleri, L., Zappavigna, V., and Ferretti, E. (2019). 'building a perfect body': control of vertebrate organogenesis by pbx-dependent regulatory networks. *Genes Dev*, 33(5-6):258–275.

Shang, Z., Yang, L., Wang, Z., Tian, Y., Gao, Y., Su, Z., Guo, R., Li, W., Liu, G., Li, X., Yang, Z., Li, Z., and Zhang, Z. (2022). The transcription factor zfp503 promotes the d1 msn identity and represses the d2 msn identity. *Front Cell Dev Biol*, 10:948331.

Shen, M. W., Arbab, M., Hsu, J. Y., Worstell, D., Culbertson, S. J., Krabbe, O., Cassa, C. A., Liu, D. R., Gifford, D. K., and Sherwood, R. I. (2018). Predictable and precise template-free crispr editing of pathogenic variants. *Nature*, 563(7733):646–651.

Shen, W. F., Montgomery, J. C., Rozenfeld, S., Moskow, J. J., Lawrence, H. J., Buchberg, A. M., and Largman, C. (1997). Abdb-like hox proteins stabilize dna binding by the meis1 homeodomain proteins. *Mol Cell Biol*, 17(11):6448–58.

Shohayeb, B., Muzar, Z., and Cooper, H. M. (2021). Conservation of neural progenitor identity and the emergence of neocortical neuronal diversity. *Semin Cell Dev Biol*, 118:4–13.

Song, X., Chen, H., Shang, Z., Du, H., Li, Z., Wen, Y., Liu, G., Qi, D., You, Y., Yang, Z., Zhang, Z., and Xu, Z. (2021). Homeobox gene six3 is required for the differentiation of d2-type medium spiny neurons. *Neurosci Bull*, 37(7):985–998.

Squair, J. W., Gautier, M., Kathe, C., Anderson, M. A., James, N. D., Hutson, T. H., Hudelle, R., Qaiser, T., Matson, K. J. E., Barraud, Q., Levine, A. J., La Manno, G., Skinnider, M. A., and Courtine, G. (2021). Confronting false discoveries in single-cell differential expression. *Nat Commun*, 12(1):5692.

Storm, E. E., Garel, S., Borello, U., Hebert, J. M., Martinez, S., McConnell, S. K., Martin, G. R., and Rubenstein, J. L. R. (2006). Dose-dependent functions of fgf8 in regulating telencephalic patterning centers. *Development*, 133(9):1831–44.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, 3rd, W. M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. *Cell*, 177(7):1888–1902.e21.

Stühmer, T., Anderson, S. A., Ekker, M., and Rubenstein, J. L. R. (2002). Ectopic expression of the dlx genes induces glutamic acid decarboxylase and dlx expression. *Development*, 129(1):245–52.

Su, Z., Wang, Z., Lindtner, S., Yang, L., Shang, Z., Tian, Y., Guo, R., You, Y., Zhou, W., Rubenstein, J. L., Yang, Z., and Zhang, Z. (2022). Dlx1/2-dependent expression of meis2 promotes neuronal fate determination in the mammalian striatum. *Development*, 149(4).

Su-Feher, L., Rubin, A., Silberberg, S., Catta-Preta, R., Lim, K., Zdilar, I., McGinnis, C., McKinsey, G., Rubino, T., Hawrylycz, M., Thompson, C., Gartner, Z., Puelles, L., Zeng, H., Rubenstein, J., and Nord, A. (2021). Single cell enhancer activity maps neuronal lineages in embryonic mouse basal ganglia. *bioRxiv*.

Tan, C. L., Plotkin, J. L., Venø, M. T., von Schimmelmann, M., Feinberg, P., Mann, S., Handler, A., Kjems, J., Surmeier, D. J., O'Carroll, D., Greengard, P., and Schaefer, A. (2013). Microrna-128 governs neuronal excitability and motor behavior in mice. *Science*, 342(6163):1254–8.

Teixeira, J. R., Szeto, R. A., Carvalho, V. M. A., Muotri, A. R., and Papes, F. (2021). Transcription factor 4 and its association with psychiatric disorders. *Transl Psychiatry*, 11(1):19.

Toresson, H., Mata de Urquiza, A., Fagerström, C., Perlmann, T., and Campbell, K. (1999). Retinoids are produced by glia in the lateral ganglionic eminence and regulate striatal neuron differentiation. *Development*, 126(6):1317–26.

Verheije, R., Kupchik, G. S., Isidor, B., Kroes, H. Y., Lynch, S. A., Hawkes, L., Hempel, M., Gelb, B. D., Ghoumid, J., D'Amours, G., Chandler, K., Dubourg, C., Loddo, S., Tümer, Z., Shaw-Smith, C., Nizon, M., Shevell, M., Van Hoof, E., Anyane-Yeboa, K., Cerbone, G., Clayton-Smith, J., Cogné, B., Corre, P., Corveleyn, A., De Borre, M., Hjortshøj, T. D., Fradin, M., Gewillig, M., Goldmuntz, E., Hens, G., Lemyre, E., Journel, H., Kini, U., Kortüm, F., Le Caignec, C., Novelli, A., Odent, S., Petit, F., Revah-Politi, A., Stong, N., Strom, T. M., van Binsbergen, E., DDD study, Devriendt, K., and Breckpot, J. (2019). Heterozygous loss-of-function variants of meis2 cause a triad of palatal defects, congenital heart defects, and intellectual disability. *Eur J Hum Genet*, 27(2):278–290.

Visel, A., Minovitsky, S., Dubchak, I., and Pennacchio, L. A. (2007). Vista enhancer browser–a database of tissue-specific human enhancers. *Nucleic Acids Res*, 35(Database issue):D88–92.

Visel, A., Taher, L., Girgis, H., May, D., Golonzhka, O., Hoch, R. V., McKinsey, G. L., Pattabiraman, K., Silberberg, S. N., Blow, M. J., Hansen, D. V., Nord, A. S., Akiyama, J. A., Holt, A., Hosseini, R., Phouanenavong, S., Plajzer-Frick, I., Shoukry, M., Afzal, V., Kaplan, T., Kriegstein, A. R., Rubin, E. M., Ovcharenko, I., Pennacchio, L. A., and Rubenstein, J. L. R. (2013). A high-resolution enhancer atlas of the developing telencephalon. *Cell*, 152(4):895–908.

Vogt, D., Hunt, R. F., Mandal, S., Sandberg, M., Silberberg, S. N., Nagasawa, T., Yang, Z., Baraban, S. C., and Rubenstein, J. L. R. (2014). Lhx6 directly regulates arx and cxcr7 to determine cortical interneuron fate and laminar position. *Neuron*, 82(2):350–64.

Wei, S., Du, H., Li, Z., Tao, G., Xu, Z., Song, X., Shang, Z., Su, Z., Chen, H., Wen, Y., Liu, G., You, Y., Zhang, Z., and Yang, Z. (2019). Transcription factors sp8 and sp9 regulate the development of caudal ganglionic eminence-derived cortical interneurons. *J Comp Neurol*, 527(17):2860–2874.

Whitington, T., Frith, M. C., Johnson, J., and Bailey, T. L. (2011). Inferring transcription factor complexes from chip-seq data. *Nucleic Acids Res*, 39(15):e98.

Wonders, C. P. and Anderson, S. A. (2006). The origin and specification of cortical interneurons. *Nat Rev Neurosci*, 7(9):687–96.

Zhang, B., Liu, M., Fong, C.-T., and Iqbal, M. A. (2021). Meis2 (15q14) gene deletions in siblings with mild developmental phenotypes and bifid uvula: documentation of mosaicism in an unaffected parent. *Mol Cytogenet*, 14(1):58.

Zhang, Z., Gutierrez, D., Li, X., Bidlack, F., Cao, H., Wang, J., Andrade, K., Margolis, H. C., and Amendt, B. A. (2013). The lim homeodomain transcription factor lhx6: a transcriptional repressor that interacts with pituitary homeobox 2 (pitx2) to regulate odontogenesis. *J Biol Chem*, 288(4):2485–500.

Zhao, L., Liu, Z., Levy, S. F., and Wu, S. (2018). Bartender: a fast and accurate clustering algorithm to count barcode reads. *Bioinformatics*, 34(5):739–747.

Zhao, Y., Flandin, P., Long, J. E., Cuesta, M. D., Westphal, H., and Rubenstein, J. L. R. (2008). Distinct molecular pathways for development of telencephalic interneuron subtypes revealed through analysis of lhx6 mutants. *J Comp Neurol*, 510(1):79–99.

Zolboot, N., Du, J. X., Zampa, F., and Lippi, G. (2021). Microrrnas instruct and maintain cell type diversity in the nervous system. *Front Mol Neurosci*, 14:646072.

Zug, R. (2022). Developmental disorders caused by haploinsufficiency of transcriptional regulators: a perspective based on cell fate determination. *Biol Open*, 11(1).