Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2019 September 18.

Published in final edited form as:

Cell Rep. 2019 August 20; 28(8): 2048–2063.e8. doi:10.1016/j.celrep.2019.07.022.

Genomic Resolution of DLX-Orchestrated Transcriptional Circuits Driving Development of Forebrain GABAergic Neurons

Susan Lindtner[1,9], **Rinaldo Catta-Preta**[2,9], **Hua Tian**[1], **Linda Su-Feher**[2], **James D. Price**[1,3], **Diane E. Dickel**[5], **Vanille Greiner**[4], **Shanni N. Silberberg**[1], **Gabriel L. McKinsey**[1], **Michael T. McManus**[4], **Len A. Pennacchio**[5,6,7], **Axel Visel**[5,6,8], **Alex S. Nord**[2,*], **John L.R. Rubenstein**[1,3,10,*]

[1]Nina Ireland Laboratory of Developmental Neurobiology, Department of Psychiatry, UCSF Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, CA 94158, USA

[2]Department of Neurobiology, Physiology and Behavior, and Department of Psychiatry and Behavioral Sciences, University of California, Davis, Davis, CA 95618, USA

[3]Development and Stem Cell Biology Program, University of California, San Francisco, San Francisco, CA 94158, USA

[4]Department of Microbiology and Immunology, University of California, San Francisco, San Francisco, CA 94158, USA

[5]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

[6]U.S. Department of Energy Joint Genome Institute, Walnut Creek, CA 94598, USA

[7]Comparative Biochemistry Program, University of California, Berkeley, Berkeley, CA 94720, USA

[8]School of Natural Sciences, University of California, Merced, Merced, CA 95343, USA

[9]These authors contributed equally

[10]Lead Contact

SUMMARY

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

*Correspondence: asnord@ucdavis.edu (A.S.N.), john.rubenstein@ucsf.edu (J.L.R.R.).

AUTHOR CONTRIBUTIONS

Conceptualization, A.S.N., S.L., R.C.-P., and J.L.R.R.; Methodology and Investigation, S.L. (ChIP-seq, RNA-seq, ISH, AB characterization, luciferase assays), H.T. (enhancer CRISPR-i), J.D.P. (ISH, RNA-seq), S.N.S. (preliminary histone ChIP-seq), G.L.M. (preliminary DLX2 ChIP-seq), D.E.D., L.A.P., and A.V. (enhancers in mice), R.C.-P. (bioinformatics, modeling), and L.S.-F. (bioinformatics); Software, R.C.-P. and A.S.N.; Formal Analysis, S.L., R.C.-P., A.S.N., and J.L.R.R.; Writing – Original Draft, S.L., R.C.-P., A.S.N., and J.L.R.R.; Writing – Review & Editing, S.L., R.C.-P., H.T., J.D.P., V.G., D.E.D., A.S.N., and J.L.R.R.; Funding Acquisition, J.L.R.R. and A.S.N.; Resources, V.G. and M.T.M. (unpublished dCas9KRAB mice); Supervision, J.L.R.R. and A.S.N.

DECLARATION OF INTERESTS

J.L.R.R. is co-founder, stockholder, and currently on the scientific board of Neurona, a company studying the potential therapeutic use of interneuron transplantation. The other authors declare no competing interests.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2019.07.022.

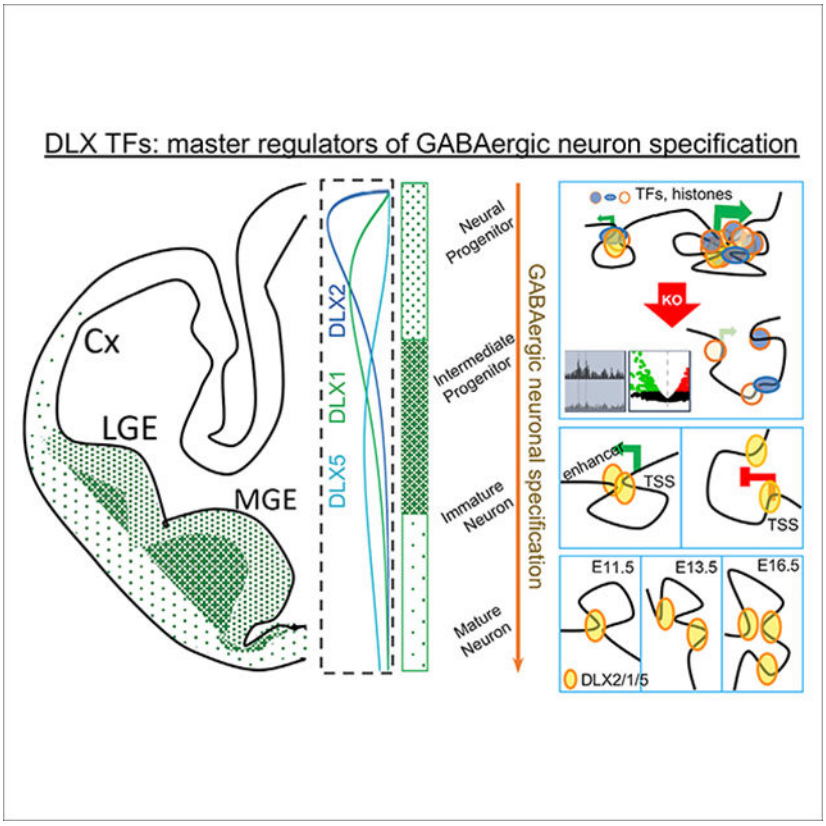
Author Manuscript Author Manuscript Author Manuscript Author Manuscript

DLX transcription factors (TFs) are master regulators of the developing vertebrate brain, driving forebrain GABAergic neuronal differentiation. Ablation of *Dlx1&2* alters expression of genes that are critical for forebrain GABAergic development. We integrated epigenomic and transcriptomic analyses, complemented with *in situ* hybridization (ISH), and *in vivo* and *in vitro* studies of regulatory element (RE) function. This revealed the DLX-organized gene regulatory network at genomic, cellular, and spatial levels in mouse embryonic basal ganglia. DLX TFs perform dual activating and repressing functions; the consequences of their binding were determined by the sequence and genomic context of target loci. Our results reveal and, in part, explain the paradox of widespread DLX binding contrasted with a limited subset of target loci that are sensitive at the epigenomic and transcriptomic level to *Dlx1&2* ablation. The regulatory properties identified here for DLX TFs suggest general mechanisms by which TFs orchestrate dynamic expression programs underlying neurodevelopment.

In Brief

Lindtner et al. reveal the regulatory wiring organized by DLX transcription factors in forebrain GABAergic neuronal specification, by integrating functional genomic, epigenomic, and genetic data on a transgenic mouse model. This network determines key sequence-encoded regulatory elements and implicates a combination of histone modifications and biophysical interactions.

Graphical Abstract



Author Manuscript Author Manuscript Author Manuscript Author Manuscript

INTRODUCTION

The development and function of all forebrain GABAergic neurons depend on a transcriptional program that is distinct from other regions of the CNS (Long et al., 2007; Silberberg et al., 2016). Near the top of this transcriptional circuitry lie the DLX homeodomain transcription factors (TFs) (Long et al., 2007). Four of the six *Dlx* genes encoded by the mammalian genome (*Dlx1, Dlx2, Dlx5,* and *Dlx6*) are expressed in the embryo by highly specific progenitor domains in the telencephalon, hypothalamus, and prethalamus (Eisenstat et al., 1999; Liu et al., 1997; Panganiban and Rubenstein, 2002). Within the embryonic telencephalon, these are the subpallial domains known at the ganglionic eminences (GEs), septum and preoptic area. The three GEs, the lateral (LGE), medial (MGE), and caudal (CGE), are the anlage for the projection neurons of the striatum, pallidum, and parts of the amygdala (Silberberg et al., 2016), as well as the interneurons that migrate to various structures including the neocortex and hippocampus (Lim et al., 2018).

Dlx RNA and DLX protein expression in the mouse GEs begin in progenitors at embryonic day 9.5 (E9.5) and follows a temporal program of *Dlx2, Dlx1, Dlx5,* followed by *Dlx6* (Eisenstat et al., 1999; Liu et al., 1997). Their expression in the GEs overlaps in the secondary progenitor domain called the subventricular zone (SVZ), suggesting that the different DLX TFs compensate for each other. Indeed, while single *Dlx* mutants have relatively mild forebrain phenotypes that affect subsets of GABAergic neurons (Cobos et al., 2005; Qiu et al., 1995; Wang et al., 2010, 2011), $Dlx1/2^{-/-}$ and $Dlx5/6^{-/-}$ double mutants have strong forebrain phenotypes (Anderson et al., 1997a; Wang et al., 2010). Constitutive $Dlx1/2^{-/-}$ mutants have poorly differentiated basal ganglia and lack most cortical interneurons (CINs). Furthermore, $Dlx1/2^{-/-}$ mutants fail to express *Dlx5* and *Dlx6* (Anderson et al., 1997a), thus generating a *Dlx*-null state, allowing one to evaluate the transcriptome and epigenome in this “ground state” lacking all DLX proteins.

Previous studies of $Dlx1/2^{-/-}$ mutants identified GE RNA expression defects at E15.5 using array technology (Long et al., 2009b). These were coupled with RNA *in situ* hybridization (ISH) to define high-resolution analyses of specific developmental effects in distinct regions and layers of the GEs (Long et al., 2009a, 2009b). This analysis led to hypotheses about potential TF regulatory networks downstream of the DLX TFs but lacked critical information about which transcription start sites (TSSs) and distant regulatory elements (REs) (e.g., enhancers and silencers) are bound by DLX proteins during GE development.

Recent single-cell RNA-sequencing (scRNA-seq) experiments have illuminated the transcriptional changes in developing GE cells, highlighting the role of key TFs including the DLX proteins (Chen et al., 2017; Mayer et al., 2018; Mi et al., 2018). However, neither scRNA-seq nor classical genetic studies have captured mechanisms underlying transcriptional wiring. Thus, major questions remain regarding how GE cell fate is encoded at genomic, epigenomic, and transcriptomic levels by TFs such as the DLX proteins. To address this need and characterize the genomic transcriptional wiring orchestrated by DLX TFs, we integrated functional genomics and genetic approaches to dissect the role of DLXs in the developing GE. We identified DLX2, DLX1, and DLX5 binding sites and target genes across GE development and measured the responses of these loci to the absence of DLX

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

proteins using transcriptomic and epigenomic methods and *in vivo* and *in vitro* studies of DLX-mediated RE function.

RESULTS

DLX2, DLX1, and DLX5 Share Binding Properties in GEs and Regulate GABAergic Neuron Specification via Activating and Repressive Functions

Toward elucidating how DLX TFs regulate transcriptional programs in the GEs, we used chromatin immunoprecipitation sequencing (ChIP-seq) and RNA-seq to compare DLX binding, chromatin state, and transcription in the GEs from wild-type and *Dlx1/2*$^{-/-}$ embryos (Figure 1A).

For ChIP, micro-dissections including MGEs, LGEs, and CGEs were prepared from wild-type (WT) E11.5, E13.5, and E16.5 embryos to compare DLX binding patterns during GE regional patterning and neurogenesis. We generated antibodies with specificity for DLX2, DLX1, and DLX5, as verified using immunofluorescence of E13.5 GEs from WT, *Dlx2*$^{-/-}$, *Dlx1*$^{-/-}$, and *Dlx5*$^{-/-}$ mice and western blot (Figure S1). We were unable to obtain ChIP-grade DLX6-specific antibody. ChIP-seq was performed and libraries were prepared for each DLX at each time point along with input and negative controls including IgG as well as three distinct blocking peptides for DLX1, DLX2, or DLX5 ChIP-seq. These blocking peptides were the antigen used in the generation of the antibodies. Finally, we performed DLX2 ChIP-seq on WT and DLX2$^{-/-}$ at E16.5; the results showed the disappearance of DLX2 specific peaks in the knockout (KO) (data not shown). Genome-wide binding patterns of all three DLXs showed high reproducibility across biological replicates and largely overlapping ChIP-seq enrichment patterns; *Gad2* and *Lhx6* loci are shown as examples (Figure 1B).

We generated a merged set of DLX2, DLX1, and DLX5 interaction sites that represented the union of all called peaks that passed strict significance thresholds (Table S3). DLX binding was enriched near loci associated with brain development in general and with more specific process, such as GABAergic specification. Target regions include promoter elements (18.1% peaks, 3,064 peaks, 7,671 total genes) and distal elements within intronic and intergenic regions (79.4% peaks, 13,440 peaks). All three DLXs showed similar patterns of an increasing number of genomic binding sites as development progressed from E11.5 through E16.5 (Figure 1C).

Many genomic regions showed DLX interaction at lower levels of enrichment. We compared the overlap of DLX binding targets for low-, medium- (the default set), and high-affinity peaks based on ChIP-seq enrichment (Figure 1F). We defined the level of affinity (stringency) by filtering out peaks above selected p values (10^{-5}, 10^{-15}, and 10^{-40} for low, medium, and high affinities, respectively) from the composition of the merged peak set across DLX ChIP-seq datasets. Using the low-affinity threshold, DLX binding was extensive across the genome (46,143 peaks), with reduced overlap among DLX2, DLX1, and DLX5 (27% common peaks). Overlap increased to 57% for medium-affinity (n = 16,919) and to 91% for high-affinity peaks (n = 4,570). Consistent with extensive binding overlap, there was high correlation of normalized ChIP-seq signal across DLXs (pairwise r^2 values > 0.7

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

for all comparisons; Figure 1D). Hierarchical clustering identified separation between E11.5 and the two later time points, followed by branching that separated DLX2 from DLX1 and DLX5.

Comparison of peak calls and genome-wide signal suggests that differences in ChIP-seq across DLXs are largely quantitative. Figure 1E is a heatmap representation separated for DLX2, DLX1, and DLX5 showing increasing binding across development across DLX TFs. Subthreshold enrichment across DLXs is present when comparing peaks that are only called in one or two of the three DLXs, further indicating that the majority of binding targets are shared (Figure S2). Notably, we identified a minority of DLX2, DLX1, and DLX5 peaks that showed specific DLX binding, although these peaks did not show enrichment for specific annotation classes. While our findings indicate that the great majority of DLX binding profiles overlap, corroborating the observed redundancy of these TFs, binding patterns may diverge in subsets of cells or later in development.

De novo motif analysis identified a core homeobox motif as the putative shared binding DNA target for all three DLXs (Figures 1H and S3). This primary TAATTA motif was strongly enriched at the center of DLX peaks and is similar to DLX binding motifs determined via SELEX analysis (Feledy et al., 1999). Highly similar homeobox motifs were also identified, suggesting that DLX TFs tolerate binding target sequence variation. Secondary motifs were also enriched within DLX peaks, suggesting combinatorial TF binding occurs at DLX target loci.

Relationship between DLX Binding and DLX-Regulated Gene Expression

We compared RNA-seq from WT and *Dlx1/2*$^{-/-}$ GE at E13.5 (n = 4 per genotype), identifying 328 differentially expressed (DE) genes at false discovery rate (FDR) < 0.05 (Figure 1G; Table S1). Downregulated genes (n = 181) were strongly enriched for terms associated with GABAergic neuron specification, while upregulated genes (n = 147) did not show enrichment for similarly specific terms (Table S2). Both upregulated and downregulated genes were enriched for annotation terms associated with general brain development. These DE results were largely consistent with previous microarray-based profiling comparing E15.5 WT and *Dlx1/2*$^{-/-}$ GEs (Long et al., 2009b).

There was a reduced distance between *Dlx1/2*$^{-/-}$ DE genes and DLX binding sites, although only for medium- and high-affinity peaks (Figure 1I), suggesting that the observed widespread lower-affinity DLX binding across the genome may be less functionally relevant than higher-affinity targets. Further supporting increased relevance of the higher-affinity DLX targets, there was a trend of increasing evolutionary conservation for DLX-bound promoter and distal elements as affinity increased for these peak sets (Figure 1J). The shared binding patterns and target motifs between DLX TFs suggest a model of competitive binding of these factors directing transcriptional activation and repression during basal ganglia and CIN development.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Loss of DLXs Leads to Changes in Chromatin State and Transcription at a Small Subset of Target Loci via DLX Binding to Distal and Proximal Regulatory Sequences

To test whether chromatin state is dependent of DLX binding, we used histone ChIP-seq to profile four histone post-translational modifications (PTMs) associated with REs and active or repressive state (H3K4me1, H3K4me3, H3K27ac, H3K27me3) in WT and *Dlx1/2*$^{-/-}$ E13.5 GEs (Figure 2A). Using WT ChIP-seq datasets, we segmented the genome into chromatin states via ChromHMM to analyze DLX binding context. Next, we identified regions that showed differential histone PTM signal in *Dlx1/2*$^{-/-}$ GE (Table S5). ChrX showed chromosome-wide differences, most likely from imbalanced sex representation in the pooled tissues used to generate these datasets. We thus restricted WT versus *Dlx1/2*$^{-/-}$ chromatin state genome-wide analysis to autosomes. Loci with significant changes in histone PTM signal in *Dlx1/2*$^{-/-}$ mutants were generally located nearby DLX peaks, suggesting that histone PTM and chromatin state changes in *Dlx1/2*$^{-/-}$ mutant GEs are directly associated with DLX binding.

In contrast to the widespread DLX binding across the genome, few DLX TF target loci (6.2%) exhibited significant changes in histone PTM signal in *Dlx1/2*$^{-/-}$ mutants (Figure S5). The finding that chromatin is unchanged in *Dlx1/2*$^{-/-}$ mutants for the majority of DLX TF binding parallels results from the intersection between DE genes and DLX TF binding, where most genes near binding sites do not show altered expression. These results suggest that, somewhat surprisingly, the majority of DLX TF binding events do not appear required for normal chromatin state or gene expression at E13.5.

We identified 510 DLX binding sites that were located within regions exhibiting loss of activating (H3K27ac, H3K4me3) and or gain of repressing (H3K27me3) signature and 574 sites with chromatin changes in the opposite direction (Table S3). We define the loci where DLX binding was required for establishment or maintenance of active transcription state as activating REs (a.REs) and the loci where DLX binding was associated with interfering with transcription as repressive REs (r.REs). As examples, the *Slc32a1* (*Vgat*) and *Otp* loci feature putative REs (a.RE and r.RE, respectively) where DLX binding was present and show significant changes to histone PTM and gene expression in *Dlx1/2*$^{-/-}$ mutants (Figure 2B). Histone PTM signal averaged across a.RE and r.RE loci changed in all four histone PTMs (Figures 2C and S5). The nearest gene to a.RE and r.RE sites was more likely to show downregulation and upregulation, respectively, in the *Dlx1/2*$^{-/-}$ mutants compared to either all DLX binding sites or high-affinity sites (Figure 2D). Most genes with strong expression changes in *Dlx1/2*$^{-/-}$ mutants were located nearby one or multiple a.REs or r.REs, as represented by red or blue circles, respectively, overlaid on the volcano plot of all DE genes, with circle size indicating significance (likelihood ratio) of differential histone signal (Figure 2E).

There were differences in function of genes associated with a.REs and r.REs (Figure 2F). a.REs were enriched for genes associated with neuronal maturation and function. In contrast, r.REs were enriched near genes associated with brain patterning, neuron fate commitment, and TF and chromatin remodeler activity. Among r.RE targets, there was a strong enrichment for other TFs (e.g., *Otp, Gsx1, Pax7*). In comparison, a.RE targets were a combination of TFs (e.g., *Dlx5, Sp8*) and lineage-specific genes (e.g., *Gad2, Slc32a1,* and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Nrxn3). This suggests that DLX TFs activate and repress coherent gene sets representing the gene regulatory network required for transitioning from progenitor to post-mitotic neurons and for maturation of GABAergic cell types.

Genes associated with a.RE and r.RE loci were additionally enriched for disease ontology terms, with increased representation of autism among a.RE and glioma and intellectual disability for r.RE loci (Figure S6). a.REs and r.REs are additionally enriched for genes linked to autism via patient mutations (p values of 0.007 and 0.03, respectively, based on permutation test), and 25/99 (25%) of high-confidence autism spectrum disorder (ASD) risk gene TSSs are within 100 kb of an a.RE or r.RE (Figure S6; Table S10).

Validation of a.RE-Mediated Activation of *Nrxn3* by DLX TFs

Nrxn3 mRNA expression was decreased in *Dlx1/2*$^{-/-}$ mutants, which was accompanied by change in local chromatin state at *Nrxn3* putative regulatory regions bound by DLX TFs (Figure 3A). To verify our regulatory model of DLX-mediated activation or repression via binding to REs, we used *in vivo* and *in vitro* approaches to assess changes in *Nrxn3* expression and DLX -dependent a.RE activity.

First, we performed *Nrxn3* RNA ISH on E13.5 WT and *Dlx1/2*$^{-/-}$ mutants (Figure 3B). *Nrxn3* expression in the SVZ of the GEs was strongly reduced in *Dlx1/2*$^{-/-}$ mutant. Next, we assessed the *in vivo* activity of an intragenic *Nrxn3* element we identified (*mm1203;* 1,450 bp) that had robust DLX binding and properties of a DLX-dependent a.RE. For the *mm1203* locus, there was H3K27ac and H3K27me3 enrichment in WT GE, and reduced H3K27ac and increased H3K27me3 in *Dlx1/2*$^{-/-}$ mutants. In a transgenic mouse assay, *mm1203* (driving *LacZ* expression) had robust and reproducible (7/7) activity in the E12.5 telencephalon, with additional activity in the eye and a small domain in the hindbrain (Figure 3C). Tissue sections showed strong enhancer activity in the SVZ and mantle zone (MZ) of the LGE and in tangentially migrating cells (probably immature CINs), and weaker activity in the MGE.

Next, we performed luciferase transcription assays in postnatal day 19 (P19) tissue culture cells, a mouse teratocarcinoma cell line, to test whether DLX proteins could regulate the activity of *mm1203.* We generated a vector in which *mm1203* is upstream of a minimal promoter and luciferase reporter gene (pGL4.23; Invitrogen). We transfected the *mm1203* luciferase reporter vector alone or in the presence of plasmids expressing DLX2, DLX1, or DLX5 (Figure 3D). DLX2, DLX1, and DLX5 activated luciferase expression by 225-, 5-, and 6-fold, respectively, in the presence of *mm1203.*

Validation of Sequence-Encoded and Context-Dependent DLX Regulation of *Arx*

After verifying DLX-mediated activation of *Nrxn3* expression, we next sought to verify sequence- and context-dependent activating and repressing function of DLX binding. To examine this, we tested the impact of DLX co-transfection with four *Arx* enhancers. The *Arx* locus is on chromosome X, and there are a number of shared DLX binding sites in the region (Figure 3E). *Arx* exhibits both cortical ventricular zone (VZ) and GE SVZ and neuronal expression in the developing telencephalon. There are four ultraconserved REs that regulate *Arx* expression (Colasante et al., 2008; Dickel et al., 2018). In transgenic mouse

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

assays, two of these *Arx* REs drove reporter expression in the GEs (*hs119* and *hs121*), whereas the other two were specific to the developing cortex (*hs122* and *hs123*).

Only the GE enhancers exhibited decreased H3K27ac in *Dlx1/2*$^{-/-}$ mutants. This suggests that while the DLX TFs bind to both the GE- and cortex-specific REs, DLX binding only had an activating effect on the GE-specific REs. When tested in the luciferase co-transfection assays, *hs119* was strongly activated by DLX2, DLX1, and DLX5 (290-, 94-, and 42-fold, respectively). *hs121* showed specific activation by DLX2 and DLX1 (57- and 7-fold, respectively). On the contrary, neither of the cortical REs, *hs122* and *hs123,* were activated by the three DLX TFs. These findings provide evidence that DLX-dependent activating and repressing activity in the GE is encoded by the primary DNA sequence of a.RE and r.RE elements.

DLX-Dependent a.REs and r.REs Are Distinguished by Genomic Sequence and Context

After verifying the biological relevance of DLX-mediated a.RE and r.RE function regulating *Nrxn3* and *Arx*, we compared *Dlx1/2*$^{-/-}$ sensitive elements (a.REs and r.REs) to each other and to DLX-bound loci that do not exhibit perturbation to chromatin state in *Dlx1/2*$^{-/-}$ mutants.

First, we tested differential representation of DNA binding motifs within a.RE and r.RE regions and between those RE regions to the ones that showed no histone PTM change to identify which motifs might distinguish activating from repressing function. For this, we considered motifs identified across all DLX peaks (Figure 1D) and added additional motifs identified specifically in a.RE (E2F3, ASCL1) or r.RE (FOXA1, GSC, OTX2) peaks (Figure S3). Using a cutoff of 1.2-fold enrichment between a.RE and r.RE regions, we identified motifs that showed directional enrichment (Figure 4A). In addition to the a.RE- and r.RE-specific motifs listed above, additional motifs showed stronger enrichment in a.REs, notably the family of basic-helix-loop-helix (bHLH) sites that includes ASCL1, AP4, NEUROG2, and OLIG2 motifs. While strongly enriched in both a.RE and r.RE peaks, there was additional increased enrichment of the primary DLX binding motif (TAATTA) in a.RE regions. Motifs that showed distinct a.RE versus r.RE enrichment were also enriched in a.RE or r.RE sequences compared to DLX-bound regions without histone signature changes.

We tested for features associated with genomic and epigenomic context of the DLX-bound loci that differentiate a.RE and r.RE elements. First, we tested for differences between promoter and distal binding patterns and across local chromatin states as identified by ChromHMM for all DLX peaks, high-affinity peaks, and a.RE and r.RE loci (Figure 4B). a.REs were enriched at distal regions (89%) and within regions with strong H3K27ac. In contrast, r.REs were skewed toward promoters (49%), and specifically toward bivalent (H3K4me3+, H3K27me3+) and active (H3K4me3+, H3K27me3−) promoters.

We assessed whether a.RE and r.RE loci and DE gene promoters showed increased local density of DLX peaks (Figure 4C). We found increased density of DLX peaks within 50 kb of DE gene promoters compared to randomly sampled matched gene sets (all gene mean, 1.7; DE mean, 2.8). Both a.RE and r.RE loci had increased density of local DLX peaks (Figure S5; Table S6). We tested whether a.RE and r.RE were more likely to be associated

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

with higher-affinity DLX binding. a.RE sites were enriched in the top quintile of DLX peaks, while r.RE loci did not show association with DLX peak strength (Figure S5). Finally, we examined correlation between a.RE and r.RE loci and histone PTM signal in WT GE, finding that a.RE and r.RE were enriched among regions with stronger histone PTM enrichment for H3K27ac, H3K27me3, and H3K4me3, but not for H3K4me1 (Figures 4D and S6).

We used logistic regression to evaluate the value of the sequence and context factors identified here to discriminate a.RE and r.RE loci (Figure 4E). For both a.RE and r.RE, chromatin context features were the strongest contributor to prediction accuracy, while DLX peak affinity and local DLX peak density were strong predictors for a.RE but not r.RE. DNA sequence motifs contributed to both models, although had less independent predictive power. Thus, sequence and context features of loci bound by DLXs were associated with the likelihood of sensitivity of these loci to DLX loss, enabling prediction of which DLX binding events have functional impact on chromatin state and gene expression in E13.5 GE without data from *Dlx1/2*$^{-/-}$ samples.

Interaction between Multiple DLX-Responsive a.REs and the *Sp8* Promoter

We identified multiple loci featuring DLX binding in a >400-kb gene desert upstream of the *Sp8* locus (Figure 5A). *Sp8* is among the most significantly downregulated genes in *Dlx1/2*$^{-/-}$ mutants (logFC = −2.4; p = 2.24E-51), with ISH analysis showing reduction in the SVZ and MZ (Long et al., 2007, 2009b). Three of the DLX-bound loci in this region overlap with previously characterized GE enhancers (*hs110, hs1007,* and *hs1226*) (Figure 5B). *hs110* and *hs1007* are among the most statistically significant DLX-responsive a.REs, with neartotal loss of H3K27ac in *Dlx1/2*$^{-/-}$ mutants. H3K27ac in the region surrounding *Sp8* also decreased in the *Dlx1/2*$^{-/-}$ mutant (Figure 5A). In contrast, for *hs1226,* only DLX2 had significant ChIP-seq signal and this region did not show H3K27ac changes in the *Dlx1/2*$^{-/-}$ mutant.

We tested the ability of these three REs to regulate *Sp8* transcription in the LGE using CRISPR interference (CRISPR-i). We grew primary cultures from E13.5 dorsal LGE from mice expressing dCas9KRAB (V.G., M. Lebedinskaya, R. Wagner, R. Jaafar, M.T.M., unpublished data). When in proximity to a TSS, dCas9KRAB downregulates transcription through steric hindrance, histone methylation, and deacetylation (Gilbert et al., 2014; Rosenbluh et al., 2017;Thakore et al., 2015). Cultured cells were infected with a lentivirus encoding a guide RNA that targeted a specific RE or the *Sp8* promoter. As a negative control we targeted a guide RNA (gRNA) to a limb RE in the same region (*hs1148*) that was weakly bound by DLX2. Transduction efficiency was >90%. 48 h later, we analyzed *Sp8* RNA levels using qRT-PCR.

gRNA targeting *Sp8* promoter reduced *Sp8* RNA up to 6-fold (Figure 5C). All three gRNAs targeting GE enhancers also reduced *Sp8* RNA. gRNA for *hs110* (the highest-affinity DLX peak) downregulated *Sp8* expression up to 4-fold, while gRNA for *hs1007* and *hs1226* reduced *Sp8* expression by ~3-fold. gRNA for the limb enhancer (*hs1148*) did not change *Sp8* RNA levels, providing evidence of gRNA/CRIPSR-i-induced *Sp8* downregulation in LGE cells that was specific to the DLX-bound GE enhancers (*hs110, hs1007,* and *hs1226*).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

To directly test whether DLXs positively modulate gene expression via binding to *hs110, hs1007,* and *hs1226,* we used the luciferase transcription assay (Figure 5D). DLX2 activated *hs110-, hs1007-,* and *hs1226*-dependent luciferase expression 50-, 6-, and 4-fold, respectively. DLX1 and DLX5 also activated *hs110* and *hs1007,* although to a lower extent than DLX2. Neither DLX1 nor DLX5 activated *hs1226,* correlating with the lack of DLX1 or DLX5 binding at this enhancer.

These results provide evidence that multiple DLX binding sites (*hs110, hs1007,* and *hs1226*) contribute to activation of *Sp8* via DLX-mediated activation and physical interaction with the promoter. These findings are consistent with the model that DLX binding is critical for establishing robust transcriptional regulation for a.RE-associated loci.

DLX TFs Mediate Spatial Activation and Repression of Core Regulatory Targets

We used ISH to examine spatial changes in expression of genes perturbed in E13.5 GEs in *Dlx1/2*$^{-/-}$ mutants, testing 6 candidates for this study; results from 2 downregulated (*Arl4D* and *Grik3*) and 2 upregulated genes (*Tox* and *Lhx9*) are shown in Figure 6A. For all 6 targets (including those in Figure S7), ISH validated predicted change in expression in *Dlx1/2*$^{-/-}$ GE. Combining our results with published ISH analyses of *Dlx1/2*$^{-/-}$ mice at E15.5 (Long et al., 2009a, 2009b) generated a set of 50 predicted downregulated and 35 predicted upregulated genes. We scored each gene for increased or decreased expression in the VZ, SVZ, and MZ to test for consistent patterns of spatial change (Figure 6B). Genes with increased RNA in the *Dlx1/2*$^{-/-}$ mutants were consistently upregulated in ISH (29, 83%), with expression increased across GE layers, but most frequently in SVZ. Of the upregulated genes, 5 (*Otp, Gsh1, Ebf3, Gbx2,* and *Pax7*) were not expressed in WT GEs, and thus were ectopically expressed in absence of DLX TFs. No downregulated genes showed ectopic expression. Downregulated DE genes were validated with reduced expression via ISH (43 genes, 86%), but a different spatial pattern of strongest decrease in the MZ followed by SVZ. These results confirm RNA-seq changes and are consistent with DLX TFs' role in controlling expression along the spatial axis of differentiation.

We next examined results from transgenic RE assays to characterize a.RE and r.RE activity as described above. We combined REs from the VISTA enhancer database (Visel et al., 2013) and two recent studies (Pattabiraman et al., 2014; Silberberg et al., 2016). Across all tested enhancers, both a.RE (24/30, 80%) and r.RE (31/39, 79%) had high rates of *in vivo* enhancer activity at E11.5. Spatial GE expression patterns in the brain were available from brain sections generated for 5 a.REs and 11 r.REs (Figures 6C, 6D, and S7). For these DLX-bound enhancers, a.REs and r.REs showed characteristic differences in activity in the telencephalon. a.RE activity in the GEs was restricted to SVZ and MZ, whereas r.RE activity in the GEs included the VZ (Figure 6D). Thus, changes in spatial expression of DLX TF-regulated genes were correlated with activity of DLX-mediated a.RE and r.RE loci.

DISCUSSION

The DLX TFs are master regulators of forebrain GABAergic neuronal development (Anderson et al., 1997a, 1997b; Cobos et al., 2007; Long et al., 2009a, 2009b; Pla et al., 2018; Yun et al., 2002). Our results show how DLX TFs organize a master regulatory

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

program by modulation of chromatin and transcriptional state, activating proneural transcriptional programs and repressing proliferative and alternative fate transcriptional programs.

Our integrated approach suggests a model for DLX function and enable construction of gene regulatory networks organized by DLX TFs in the GE (Figure 7A). As *Dlx2/1/5* are co-expressed in individual cells, data indicate that they compete for binding to homeobox DNA motifs at shared genomic targets that expanded across development. DLX TF binding could drive either transcriptional activation (largely via interactions at distal elements) or repression (largely via binding near the TSS), while most DLX binding sites did not show a large effect with regard to changes in chromatin or transcriptional state in $Dlx1/2^{-/-}$ GEs. In contrast, a subset of loci showed chromatin and transcriptional states that were highly sensitive to the $Dlx1/2^{-/-}$ state.

Our analysis highlights emergent features of regulatory DNA sequence and chromatin context that in part explain variability among individual DLX binding targets in sensitivity to DLX loss. Using *in vivo* and *in vitro* approaches, we show that the sensitivity of REs to DLX-mediated activation or repression (i.e., a.RE and r.RE loci) is encoded by their DNA sequence. TF motifs associated with a.REs (e.g., ASCL1 and OLIG2) and r.REs (e.g., OTX2) are TFs with major functions in GE biology (Casarosa et al., 1999; Hoch et al., 2015; Long et al., 2009b; Petryniak et al., 2007). This is consistent with a model of combinatorial and/or competitive binding by TFs and co-regulators.

The transgenic RE and cell culture transcription assays performed here recapitulate our predicted regulatory activity and sensitivity to DLXs for *Nrxn3, Arx,* and *Sp8* loci. In parallel, genomic context of DLX binding sites matters, as local chromatin state and inferred biophysical interactions were among the strongest predictors of epigenomic and transcriptional sensitivity to DLX loss. Specifically, a.RE and r.RE elements were associated with loci that featured clustered DLX binding and high rates of activating and/or repressing histone PTMs.

We propose that principles elucidated for the DLXs represent common characteristics of TFs that organize cellular differentiation and metazoan development. Predicting the impact of regulatory sequence variation, for example in evolution and disease, can be improved via statistical prioritization of TF binding events that are likely to have strong functional relevance. Previous studies found similar disagreement between the binding targets of a TF and transcriptional sensitivity of these targets to loss of the TF (Sandberg et al., 2016; Cusanovich et al., 2014), but they were not able to dissect the underlying nature of this disconnect. Our data are in line with the idea that enhancer-promoter and DNA-RNA-protein interactions underlie robust transcriptional regulation and phase separation for these critical genes (Boija et al., 2018; Hnisz et al., 2017). Our results suggest that the sequence-encoded RE function, in tandem with biophysical interactions, may explain how master TF binding establishes developmental transcription patterns.

The genes that are directly regulated by DLX-mediated a.RE and r.RE activity are strongly enriched for TFs, lineage-defining molecules, and disease-relevant loci. DLX regulatory

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

activity can be separated by DLX2/1/5 activity in the VZ, SVZ, and MZ. Figure 7B highlights the DLX TF gene regulatory network of TFs, non-TFs, and genes relevant to brain disorders. All of the genes depicted have DLX binding at the TSS or within the genomic region. For the majority of gene loci (45/71 TFs and 18/35 non-TFs), their RNA expression is altered (ISH and/or RNA-seq), and gene-associated RE epigenetic states are changed (a.RE or r.RE) in the *Dlx1/2*$^{-/-}$. Below we discuss specific developmental programs that are orchestrated by DLX-dependent transcriptional circuitry.

DLX proteins are nearly ubiquitously expressed in the SVZ (SVZ1 and SVZ2; Eisenstat et al., 1999; Petryniak et al., 2007), where we propose they have a core function in promoting specification of forebrain GABAergic neurons. DLX2 and DLX1 promote activation of *Dlx5* and *Dlx6* (Anderson et al., 1997b), thereby generating cells with four degrees of DLX redundancy, creating an evolutionarily buffered foundation for ensuring the generation of forebrain GABAergic neurons.

DLX2, and subsequently DLX1, are co-expressed in subsets of VZ and most SVZ progenitor cells, where they promote proneural programs (Eisenstat et al., 1999; Yun et al., 2002). These include transcriptional repression of TF genes that regulate progenitor cell states, including *Gsx2, Hes5, Pax6, Olig2,* and *Otx2.* Some of these TFs also regulate forebrain regional fate (e.g. *Gsx2, Otx2,* and *Pax6* [Hoch et al., 2015; Toresson and Campbell, 2001; Yun et al., 2001]) or promote oligodendrogenesis (*Olig2* [Petryniak et al., 2007]). DLXs also promote proneural programs through repressing *Notch1* and *Delta1* (*Dll1*). In the absence of DLX proteins, the GEs have an ectopic expression of markers of the hypothalamus (*Otp*), midbrain (*Pax7*), pallium (*Ebf3, Lhx9*), as well as LGE ectopic expression of MGE genes (*Gbx1, Gbx2*) (Long et al., 2009b).

There is evidence for additional transcriptional pathways that drive GE specification and neurogenesis, which are driven by ASCL1 and GSX2. *Ascl1*;*Dlx1/2* triple mutants lose characteristic GE molecular features (Long et al., 2009a; 2009b). Furthermore, removal of *Gsx2* in the *Dlx1/2* mutants rescues overexpression of *Ascl1* (Wang et al., 2013). Neither the *Ascl1* or *Gsx2* locus show RE epigenetic changes (Figure 7B), suggesting that they function primarily upstream or parallel to *Dlx* genes.

DLX-driven transcription generates properties fundamental to GE cells, and subsets of their derived neurons, including expression of *Gad1, Gad2, Slc32a1* (GABA regulatory genes), and other non-TFs including *Cxcr4, Cxcr7* (*Ackr3*), *Erbb4, Nrnx3,* and *Robo2.* DLX activates GE expression of multiple TFs including *Arx, Lhx6, Lhx8, Maf, Mafb, Sp8, Sp9,* and *Zeb2* (*Zfhx1b*), most with distinct functions in specific neuronal lineage. For instance, *Maf* and *Mafb* regulate development of MGE-derived cortical interneurons; together, they repress the generation of somatostatin^{+} interneurons (Pai et al., 2019) *Sp8&9* specify the identity and differentiation of striatopallidal projection neurons (Xu et al., 2018). DLX TFs additionally repress inhibitors of neuronal differentiation/maturation (e.g., *Id2, Id4,* and *Ascl1*).

We provide evidence for DLX-driven expression of neuronal features in the MZ. These include repression of hypothalamic genes (*Hmx3/Nkx5–1, Otp*), activation of striatal and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

pallidal projection neuron genes (*Gnal, Gucy1a3, Npas1, Rxr*) and interneuron genes (*Cxcr4, Cxcr7, ErbB4, Th*). The Gene Ontology (GO) terms for genes dysregulated in *Dlx1/2*$^{-/-}$ (Figure 2F) include axon, synaptic transmission, and neuron apoptotic process; these provide evidence that DLX, through a.REs and r.REs, regulate multiple processes in maturing neurons. Indeed, *Dlx* mutants have defects in neuritic processes, synapse function, and neuronal survival (Cobos et al., 2007; Pla et al., 2018).

In addition to coordinating specific developmental programs, DLX-dependent transcriptional circuitry regulates many genes that predispose to neurodevelopmental disorders. For example, *Arx, Maf,* and *Zfhx1b* are targets of DLXs; mutations in these genes cause human intellectual disability that is linked to abnormal CINs in mouse mutants (Mowat et al., 2003; Niceta et al., 2015; Olivetti and Noebels, 2012). Defining DLX-dependent transcriptional circuitry provides a framework for understanding disease-causing pathways.

Our results demonstrate that while TF ChIP-seq provides critical interaction maps, chromatin interaction alone is not sufficient to elucidate the functional relevance of binding. By identifying the sequence and context constraints and showing that sensitivity to DLX TF ablation can be modeled without incorporating null mutant data, this study provides a path forward to move from TF binding catalogs, to a functional understanding of the gene regulatory wiring driving brain development.

STAR★METHODS

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources should be directed to and will be addressed by the Lead Contact, John L. R. Rubenstein (John.Rubenstein@ucsf.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice and genotyping—All procedures and animal care were approved and performed in accordance with National Institutes of Health and the University of California San Francisco Laboratory Animal Research Center (LARC) guidelines. *Dlx1*$^{+/-}$, *Dlx2*$^{+/-}$ and *Dlx1/2*$^{+/-}$ mice were genotyped as in Qiu et al. (1995); *Dlx5*$^{+/-}$ mice were genotyped as in Depew et al. (1999). For RNA-seq experiments, an equal number of males and females were used. ChIP-Seq and native histone ChIP-Seq was performed on *Mus musculus* CD1 strain at developmental stage E11.5, E13.5 or E16.5. The embryos were not assessed genotypically for gender since we used a pool of embryos and therefore expect a roughly equal number of male and females.

Transgenic enhancer assays—Transgenic enhancer assays were performed at Lawrence Berkeley National Laboratory (LBNL) under the approval of the Animal Welfare and Research Committee (AWRC). Mice for transgenic assays were housed at the Animal Care Facility (ACF) of LBNL, monitored daily for food and water intake, and inspected weekly by the Chair of the AWRC and the head of the animal facility in consultation with the veterinary staff. The LBNL ACF is accredited by the American Association for the Accreditation of Laboratory Animal Care (AAALAC). Transgenic assays were performed in the *Mus musculus* FVB strain at developmental stage E11.5 or E12.5. The resulting embryos

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

were not assessed phenotypically for gender, which is not outwardly obvious at these ages. Therefore, we expect that a roughly equal number of male and female embryos were assessed.

METHOD DETAILS

DLX antibody production—Antibodies against DLX1, DLX2 and DLX5 were made in rabbits by GenScript. Antibodies against DLX2 were raised against amino acids (aa) 1–154 and N-terminally tagged with a Maltose binding protein (MBP) (Kuwajima et al., 2006). Antibodies against DLX1 were raised against aa 1–121 and N-terminally tagged with His-Tag. Antibodies against DLX5 were raised against aa 188–289, N-terminally tagged with His. Antibodies were purified using either high affinity M or lodoacetyl resin, based on whether the protein had the MBP- or His Tag, respectively. To assess the specificity of the antibodies, Western Blots were performed using E13.5 or E16.5 basal ganglia nuclear extract from either WT or the respective constitutive *Dlx* mutants (Figure S1A). Furthermore, the antibodies against DLX2/1 and 5 were also tested using immunohistochemistry assays of E13.5 telencephalon sections of WT and *Dlx2*$^{-/-}$, *Dlx1* or *Dlx5*$^{-/-}$ respectively (Figure S1B).

Chromatin immunoprecipitation (ChIP)—ChIP was performed using the DLX1, DLX2 and DLX5 antibodies described above. Basal ganglia were dissected in cold PBS from CD1 embryos (6 litters/Ab for all the DLX ChIP-seq at E11.5; 2 litters/Ab for the DLX2 ChIP E13.5; 3 litters/Ab for the DLX1 and DLX5 ChIPs E13.5; and 1 litter/Ab at E16.5). The basal ganglia consisted of the LGE, MGE and CGE progenitor and mantle zones; they were fixed in 1.5% formaldehyde at RT for 20 min, neutralized with glycine and washed gently in PBS. The fixed cells were lysed with a hypotonic buffer (50 mM Tris pH 7.5 / 0.5% NP40 / 0.25% Sodium Deoxycholate / 0.1% SDS / 150 mM NaCl) to obtain the nuclei; these were then lysed in 1% SDS buffer and the chromatin was sheared into 300–1000 bp fragments by sonicating for 40 cycles (30 s on and 45 s off) using a bioruptor (Diagenode). Immunoprecipitation (IP) reactions were performed with the sheared chromatin diluted 1/10 times with “dilution buffer” (0.01% SDS, 1.1% Triton X-100, 1.2mM EDTA, 16.7mM Tris-HCl, pH 8.1, 167mM NaCl, usually in 6 ml. Antibody was then added: 5 μg DLX specific Ab. Negative control CHIP reactions used either rabbit IgG (5μg) or blocking peptide (DLX antigen used for immunizing rabbits; 50x molar excess). Antibody/chromatin complexes were purified using Dynabeads (Invitrogen) and washed extensively in 4 different “wash buffers” (low salt (1st), 0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl, pH 8.1, 150mM NaCl; high salt (2nd), 0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl, pH 8.1,500mM NaClLiCl (3rd), 0.25M LiCl, 1% IGEPAL CA630, 1% deoxycholic acid (sodium salt), 1mM EDTA, 10mM Tris, pH 8.1and TE, 4th).

Complexes were eluted in 1% SDS, 10mM Sodium bicarbonate buffer at 65°C for 10 min. Eluted chromatin was reverse cross-linked overnight at 65°C in the presence of 500mM NaCl, then subsequently treated with RNase (10 μg/ 200 μl reaction, 15 min at 37C) and Proteinase K (10 μg/ 200 μl reaction, 60 min at 55C) and cleaned using a ChIP DNA Clean & Concentrator kit (Zymo Research). The chromatin was quality controlled (QC) using

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

qPCR to check for enrichment of genomic DNA fragments that were expected, and not-expected, to have DLX binding.

Libraries were prepared using an Ovation Ultralow DR Multiplex System (Nugen), size selected in the range of 300 bp on a chip from BluePippin (Sage Science) and lastly QC tested on a Bioanalyzer (Agilent).

The libraries were sequenced as single-ended 50 nucleotide reads on a HiSeq 4000 (Illumina) at Center for Advanced Technology (UCSF).

Native histone ChIP—Each native histone ChIP was performed starting with ~250,000 nuclei from WT and *Dlx1/2*$^{-/-}$E13.5 basal ganglia. The native ChIP was performed as described earlier described (Magklara et al., 2011). Briefly, nuclei were extracted and digested with micrococcal nuclease (MNase, Sigma). A population of mono- and di-nucleosomes were used in chromatin immunoprecipitation assays. Antibodies used were specific to H3 monomethyl lysine-4 (H3K4me1, Abcam, ab8895), H3 trimethyl lysine-4 (H3K4me3, Abcam, ab8580), H3 trimethyl lysine-27 (H3K27me3, Active Motif, 39157) and H3 acetylated lysine 27 (H3K27ac, Abcam, ab472). Immuno-precipitated DNA was washed, isolated and cleaned as for the TF ChIP-Seq described above.

gRNA design and lentivirus infection of primary dorsal LGE dCas9-KRAB culture—The *dCas9-KRAB* mice were generated in the FVB background with the TARGATT™ site-specific knock-in technology (Tasic et al., 2011) by introducing into the *Hipp11* locus a construct expressing from the CAG promoter: puromycin resistance, mCherry and the dead Cas9 (dCas9) protein fused to the KRAB (Krüppel Associated Box) domain. The dcas9-KRAB protein lacks endonuclease activity but retains its ability to bind to single guide RNAs (sgRNAs) for specific gene targeting and silencing through heterochromatin formation induced by the KRAB domain (V.G., M. Lebedinskaya, R. Wagner, R. Jaafar, M.T.M., unpublished data). Guide RNA oligonucleotides were annealed and cloned into *U6-stuffer-longTracer-GFP* lentivirus vector. gRNA lentivirus were produced in 293T cells through transfection of packaging plasmids psPax2, pmD2G (Addgene) and gRNA lentivirus vector. Concentrated virus was used to infect primary dorsal LGE culture deriving from *dCas9-KRAB* mice at the time of seeding. The RNA from the culture was purified 48 hr later and assayed for target gene expression via RT-QPCR. List of gRNAs used:

SP8TSS: TTGGAGCTACAATTGTGGCGGAAT

AAACATTCCGCCACAATTGTAGCT

hs110: TTGG CTTTGTACTGCGGCTCAATT

AAACAATTGAGCCGCAGTACAAAG

hs1007: TTGG GTAACAGATGTTCGCTAATC

AAAC GATTAGCGAACATCTGTTAC

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

hs1226: TTGG GGCCGACTTATGACAGCCTC

AAACGAGGCTGTCATAAGTCGGCC

hs1148: TTGG CCGGATCCAGCCGTTTCCAC

AAACGTGGAAACGGCTGGATCCGG

qPCR and primers—We used a SYBR green mix (PerfeCTa SYBR Green FastMix ROX, Quanta) including, buffer, polymerase, dNTPs and SYBR green for the qPCR reactions. We ran 10 µl reactions on a 7900HT Fast Real-Time PCR system (Applied Biosystems). We always used negative control primer-sets for the calculated $2^{-\Delta\Delta CT}$ value and analyzed as in Vokes et al. (2007). For DLX ChIP-seq QC we used the following primers:

DlxI12b: F: CGGGCCCATCAAACACAAC

R: TGGGCGAAAAAATTGCTCAT

DlxI5/6: F:GCCATCCAATTTGAAGCAG

R: GACGGTAAAAACGCTGTAATCAG

Internal control upstream of Dlx2: F:CAGGACTAAGCAGGCCTTTG

R: TGACCCCAATGACTCTCCAC

For the *dCas9-KRAB Sp8* RT-qPCR experiments we used two *Sp8* gene specific primers for first-strand cDNA synthesis

GSP1 for Sp8: CAGAGGAGTGGATCCCAACC

GSP2 for Sp8: CAGGTAGCAGCAAGCATGGC

Sp8-specific qPCR primers: F: GCGCACACTTGCACCATATC

R: TTCTTCTCGCGTTCCCCTTC

As internal control we used a gene specific primer against the *Etv1* locus for first-strand cDNA synthesis:

GSP for Etv1: TGAACATGGGCTGTGGGGTT

Etv1-specific qPCR primers: F: CATCAGGACGGGATGCTTCA

R: GAGGCCATGAAAAGCCAAACT

RNA-seq—Total RNA from four embryos (E13.5, 2 males and 2 females) of either WT or *Dlx1/2*$^{-/-}$ was purified using RNeasy Mini Kit (QIAGEN).The DNase-treated RNA was checked on Bioanalyzer and RNA-seq was performed at the UCSFSABRE Functional

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Genomics Facility. Briefly, RNA sequencing libraries were generated using TruSeq Stranded Total RNA Library Prep Kit with Ribo-Zero Mouse, according to the manufacturer's protocol (Illumina, San Diego, CA, USA). Fragment size distribution was assessed using the Bioanalyzer 2100 and the DNA high-sensitivity chip (Agilent). Concentration of the libraries was measured using the Kapa library quantification kit (Kapa Biosystems, Woburn, MA, USA). Libraries were multiplexed, sequenced on the HiSeq 4000 and generated, roughly 814 million single-end 50bp reads.

***In situ* RNA Hybridization (ISH)—**18 μm frozen cryostat sections of E13.5 mouse head were dried, wash three times with PBS (5 min each) and steamed for 10 min in Sodium Citrate buffer (10 mM, pH6). Sections were cooled down to RT for 15 min then rinsed with PBS three times (5 min each). Acetylation was performed for 10min in Acetylation Buffer containing 1.3% triethanolamine, 0.17% HCl and 0.4% acetic anhydride in water. Sections were then rinsed with PBS three times (5 min each) and prehybridized by incubating with hybridization buffer (50% formamide, 5XSSC pH 4.5, 50ug/ml yeast tRNA, 1% SDS, 50ug/ml Heparin) for 1 hour in 67C oven. After prehybridization *in situ* probes diluted in hybridization buffer at 500ng/ml were added for overnight incubation at 67C. Next day slides were rinsed with pre-warmed 5X SSC pH 4.5, washed twice (30min each) with 0.2X SSC pH4.5 at 70C and then wash once (5min) with 0.2X SSC pH4.5 at room temperature followed by a wash with NTT buffer (0.15 M NaCl, 0.1 M Tris pH 8.0 0.1% Tween-20). Sections were blocked with NTT blocking buffer containing 5% heat inactivated sheep serum and 2% Blocking buffer (Sigma-Aldrich) for 1 hour at room temperature followed by an overnight incubation at 4C with Anti-Digoxigenin-AP antibody (1:5000 dilution in NTT blocking buffer). Next day sections were washed three times with NTT buffer (30 min each), followed by three 5-min washes with NTTML buffer (0.15 M NaCl, 0.1 M Tris pH 9.5, 0.1% Tween-20, 50 mM $MgCl_2$, 2 mM Levamisole) and incubated with developing reagent BM Purple (Sigma- Aldrich) until desired intensity of the signal was reached. Development reaction was stopped with PBS. Sections were fixed in 4% PFA o/n and were coverslipped in Aqua-Mount (Fisher Scientific).

Primers used to generate *in situ* hybridization (ISH) probes

Arl4d: F: CTTGACTGAAATGGCCCCTA

R: ACCGTCTCTTCTTGCTCGAC

Grik3: F: CACCAACCCCAGTGTCTTCT

R: GTTGATAACCGCATCCGTCT

Tox: F: AGATTGGGAACGCAGAAGGA

R: AGTGTAAGGCAGTCTGGACC

Lhx9: Rétaux et al., 1999

Nrxn3: Püschel and Betz, 1995

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Pax7: Jostes et al., 1990

Zbtb20: F: CAGGTAATCCTGGCTCATGC

R: GGCACCTACCCTGTGTTCTC

Immunofluorescent tissue staining—Immunofluorescent labeling was performed on 18 μm (E13.5) cryosections with the following primary antibodies: rabbit anti-DLX1, rabbit anti-DLX2, rabbit anti-DLX5 (production of these described above). 546 Alexa-conjugated secondary antibodies (Life Technologies) were used. Sections were coverslipped with Aqua-Mount.

Luciferase reporter assay in P19—RE elements were cloned into a *pGL4.23-Luciferase reporter vector* (Promega) at the Xhol and BglII sites. P19 cells (ATCC) were plated in 96 well plates, transfected with a total of 200 ng DNA using X-tremeGENE HP DNA transfection reagent (Sigma-Aldrich). Luciferase assays were performed 48h after transfection using Dual Luciferase Reporter Assay System (Promega). A Renilla reporter was used for normalization. Statistical differences between experimental groups were determined with paired t test.

List of primers used to clone RE elements into pGL4.23

mm1203: F: GTGCCAACAGTGTCTAGATTA

R: TCTACAGGCATTCCCAAAGC

hs119: F: GAATTTGAGCTGAAAACATTTCC

R: TTGGATTCTGGGAAGAATCC

hs121: F: GCGAGGCGATGTAAAGAATG

R: TTGAGTAGACAAACATGTTTTC

hs122: F: GATGCAAATACTTCTAGAATGG

R: TAAAGTCAATCCCATGATAGTT

hs123: F: GCCCTATCTTATGGAAGGAAC

R: TGCCACATGGAGATGAAGAG

hs1007: F: GCCTGTTTGTACATTAGCGAA

R: TCAACAACAACAAAAAGACTAAA

hs110: F: ACCCATATTTTGTTGAGCT

R: TGTAGACATCTGATGGCA

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

hs1226: F: CTGGAAAGGCCACC

R: CAGGTCTTCGCCTGA

Western blot analysis—Nuclear preparation from basal ganglia tissue from one litter of either E13.5 or E16.5 MGE was prepared using a NE-PER kit (ThermoFisher) with protease inhibitor (Sigma-Aldrich) and the total protein was separated in SDS-PAGE gels and transferred to nitrocellulose membranes. Western blotting was performed by standard procedures.

Transgenic enhancer assays—Enhancer names (mm/hs numbers) are the unique identifiers used in the VISTA Enhancer Browser (https://enhancer.lbl.gov). Candidate enhancer, mm1203, was PCR-amplified and cloned into an hsp68-promoter-lacZ reporter vector (Pennacchio et al., 2006). Transgenic assays were performed according to published methods (Kothary et al., 1989; Pennacchio et al., 2006).

Briefly, the enhancer-reporter vector was linearized and injected into the pronucleus of FVB strain single cell stage mouse embryos (E0.5). Embryos were implanted into surrogate CD-1 strain *Mus musculus* mothers and were then collected and stained for reporter gene expression at E11.5 or E12.5. Embryos were excluded from analysis only if they did not harbor the transgene or if they were not at the correct developmental stage. No comparisons were made between cohorts of transgenic embryos, so randomization and experimenter blinding were unnecessary and not performed. Sample sizes were determined empirically based on our experience performing > 2,000 transgenic enhancer assays. Only *LacZ* activity patterns that were observed in three or more different embryos resulting from independent transgene integration events of the same construct were considered reproducible.

QUANTIFICATION AND STATISTICAL ANALYSIS

Differential Gene Expression—We determined the differential gene expression between the wild-type and $Dlx1/2^{-/-}$ samples by individually aligning the FASTQ reads from the RNA-seq experiment using STAR (version 2.4.2a; (Dobin et al., 2013), after QC'ing the files with FASTQC (version 0.11.6; Andrews, 2010) and counting reads with featureCounts (version 1.5.1; Liao et al., 2014). Following quality filtering and alignment, we analyzed the samples using a custom R script running edgeR (version 3.24.2) / limma (version 3.38.3; Phipson et al., 2016; Robinson et al., 2010), that used a statistical method that ascertained the differential gene expression levels.

ChIP-seq Data Analysis—FASTQ files containing the reads were QC'ed using FASTQC utility, and the remaining adapter sequences were trimmed using Trim Galore (version 0.4.5; Krueger, 2015). The resulting reads were aligned to the mouse (mm9) genome using Burrows-Wheeler Alignment (BWA; version 0.7.9a; Li and Durbin, 2009), duplicates removed with Samtools (version 1.8; Li et al., 2009). Peaks of enriched binding regions against both negative binding and input DNA control were called using MACS (version 2.1; Zhang et al., 2008) (Figure S4). Using DeepTools (version 2.5.3; Ramírez et al., 2016), we determined the Pearson correlation among the reads of all DLX dataset replicates (Figure S2). After removing the outlier replicates, the remaining ones were merged and analyzed

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

again, showing the correlation among the different DLXs at the different growth stages (Figure 1D). Narrow peaks of DLXs were overlapped and merged into one dataset using custom R scripts, where they were annotated for neighboring gene regulatory regions and filtered against repeat, blacklisted and gapped regions. Additional annotation was made for each DLX-bound peak overlapping evolutionary conserved regions, as well as VISTA transcriptional enhancers (Table S3).

Histone (H3K27ac, H3K27me3, H3K4me3, H3K4me1) broad peaks, also called using MACS2, were then overlapped with and annotated on the base peak set described above. Furthermore, MACS2 bdgdiff was used to assess the differential intensity of the histone PTMs as measured by ChIP-seq. We used the same approach to determine local H3K4me1 changes in WT versus KO, as a marker for transcriptional enhancer establishment or hinderance. We used a p value cutoff of 1E-4, with 350 as minimum length of peak for all narrow and broad peak calls; for differential histone peak calls we used a likelihood ratio cutoff of 1E+3. Command-line examples for the software used are provide in Table S7.

We used ngsplot (version 2.63; Shen et al., 2014) within custom scripts to display overall genome wide coverage, both as coverage plots and heatmaps, for a variety of stratifications (Figures 1E and 2C), with a fragment length of 150 bp. Moreover, we used DeepTools to make peak complexity plots (not shown). Peak complexity was further analyzed by assessing the number of peaks within each peak neighborhood and finding the statistical distribution of their means, both overall and within a.REs and r.REs (Figure S5); we ran a custom R script that performed permutation analysis at 95% confidence level to determine the statistical significance of those differences.

Chromatin State Inference—We inferred chromatin state change as a result of DLX binding, herein designated as a.REs and r.REs (activating and repressing regulatory elements, respectively), by the local occurrence of H3K4me3, H3K27ac, H3K27me3 signature mark changes in WT samples in relation to *Dlx1/2*$^{-/-}$ (KO) ones. Following determining significantly altered PTM regions in WT versus KO, we assigned a.REs to DLX-bound regions with increases in H3K4me3 or H3K27ac, or decreases in H3K27me3, in WT. Conversely, r.REs were assigned to regions with changes in the opposite direction; bound regions with dual assignment were disregarded. We inferred putative enhancer regions when the H3K4me1 was present in WT; the effect of DLX binding to putative enhancers was assessed by the differential likelihood of H3K4me1 in WT in comparison with *Dlx1/2*$^{-/-}$.

We evaluated the correlation among histone marks at DLX TF-bound regions, intensity of histone ChIP-seq signal in WT, peak neighborhood complexity and differential gene expression, visually (Figure 5E) and by running ANOVA on the factors against gene expression fold change (Table S8).

Furthermore, we created an 8-state HMM model using our WT histone data and ChromHMM (Ernst and Kellis, 2017) and assigned chromatin states based on the emission probabilities. We further segmented the states into proximal and distal, as appropriate for the emission signal.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Motif Analysis—*De novo* motif discovery was performed using HOMER (version 4.9; Heinz et al., 2010) with standard parameters, 250 bp up- and downstream of DLX peaks. Upon selection of significant classes of overrepresented motifs on all DLX-bound regions based on FDR corrected (Benjamini-Hochberg method) p value < 0.05 and non-redundancy, we carried out enrichment analysis of those motifs on those regions, as well as regions stratified into distal and proximal sites, and a.REs and r.REs (Figures 1H and S3). All motif enrichment determination was also made with HOMER. Command-line examples for the software used are provided in Table S7.

We further analyzed the difference between motif enrichment in a.RE versus r.RE for HOMER known motif database using custom R scripts (Figure 4A) and found statistically significant differences between a.RE- and r.RE-relative enrichment by calculating means differences using both a bootstrapping approach and an unpaired two-sample Wilcoxon test (Table S9).

Gene Regulatory Network—We constructed a core gene regulatory networks (GRN) at the three growth stages with Cytoscape (version 3.6.1; Shannon et al., 2003) by assigning DLXs as source nodes and DE genes with annotated DLX-bound peaks at presumed REs, significant expression difference and chromatin state change (a.RE/r.RE) as target nodes (not shown). Node colors were assigned to as green for activation or downregulation in *Dlx1/2*$^{-/-}$ samples and red for inhibition or upregulation in KO, with intensity proportional to gene expression fold change. Edges were also colored as green and red as above, with widths proportional to the p values of the respective gene expression fold changes.

Gene Ontology Analysis—We conducted gene ontology analysis using the GREAT algorithm for the whole peak dataset and compared with the results for the peaks showing in a.REs and r.REs, by means of a custom R script running the rGREAT package (version 1.14.0; Gu, 2015). We found a number of differentially enriched GO terms in a.RE versus r.RE and selected a few relevant ones for our test system (Figure 2F); the complete results are on Table S4. The genes that differentially made the list for relevant disease GO terms are shown in Table S10 and are graphically depicted in Figure S6.

Modeling sequence and context features for predicting a.RE and r.RE loci—Logistic regression analysis was performed with a.RE, r.RE or no change class as the dependent variable and motif presence, histone PTM peak fold enrichment, DLX2 E13.5 peak fold enrichment, number of DLX peaks within 50KBP of peak, ChromHMM state and distal versus proximal location. a.RE or r.RE were compared against all other peaks. ROC curves were generated based on predicted versus observed results for the logistic regression model and area under the curve was calculated using discrete approximation.

DATA AND CODE AVAILABILITY

The genomic and epigenomic data generated in this study and presented in this publication have been deposited in NCBI database and are accessible through GEO Series accession number GSE124936 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE124936) and can be visualized in UCSC track hubs whose information is provided on Nord Lab

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

GitHub page (https://github.com/NordNeurogenomicsLab/Publications/tree/master/Lindtner_Cell_2019).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

J.L.R.R. (Nina Ireland) was supported by NIMH grant R37/R01 MH049428. A.S.N., R.C.-P., and L.S.-F. were supported by NIGMS grant R35 GM119831. R.C.-P. was supported by a Science without Borders Fellowship from CNPq (Brazil). L.S.-F. was supported by the UC Davis Emmy Werner and Stanley Jacobsen Fellowship. Research conducted at the E.O. Lawrence Berkeley National Laboratory was supported by NIH grant R01HG003988 (to L.A.P.) and performed under U.S. Department of Energy contract DE-AC02-05CH11231, University of California.

REFERENCES

Anderson SA, Eisenstat DD, Shi L, and Rubenstein JL (1997a). Interneuron migration from basal forebrain to neocortex: dependence on Dlx genes. Science 278, 474–476. [PubMed: 9334308]

Anderson SA, Qiu M, Bulfone A, Eisenstat DD, Meneses J, Pedersen R, and Rubenstein JL (1997b). Mutations of the homeobox genes Dlx-1 and Dlx-2 disrupt the striatal subventricular zone and differentiation of late born striatal neurons. Neuron 19, 27–37. [PubMed: 9247261]

Andrews S (2010). FASTQC. A quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/.

Boija A, Klein IA, Sabari BR, Dall'Agnese A, Coffey EL, Zamudio AV, Li CH, Shrinivas K, Manteiga JC, Hannett NM, et al. (2018). Transcription factors activate genes through the phase-separation capacity of their activation domains. Cell 175, 1842–1855.e16. [PubMed: 30449618]

Casarosa S, Fode C, and Guillemot F (1999). Mash1 regulates neurogenesis in the ventral telencephalon. Development 126, 525–534. [PubMed: 9876181]

Chen Y-JJ, Friedman BA, Ha C, Durinck S, Liu J, Rubenstein JL, Se-shagiri S, and Modrusan Z (2017). Single-cell RNA sequencing identifies distinct mouse medial ganglionic eminence cell types. Sci. Rep 7, 45656. [PubMed: 28361918]

Cobos I, Calcagnotto ME, Vilaythong AJ, Thwin MT, Noebels JL, Baraban SC, and Rubenstein JLR (2005). Mice lacking Dlx1 show subtype-specific loss of interneurons, reduced inhibition and epilepsy. Nat. Neurosci 8, 1059–1068. [PubMed: 16007083]

Cobos I, Borello U, and Rubenstein JLR (2007). Dlx transcription factors promote migration through repression of axon and dendrite growth. Neuron 54, 873–888. [PubMed: 17582329]

Colasante G, Collombat P, Raimondi V, Bonanomi D, Ferrai C, Maira M, Yoshikawa K, Mansouri A, Valtorta F, Rubenstein JLR, and Broccoli V (2008). Arx is adirect target of Dlx2 and thereby contributes to the tangential migration of GABAergic interneurons. J. Neurosci 28, 10674–10686. [PubMed: 18923043]

Cusanovich DA, Pavlovic B, Pritchard JK, and Gilad Y (2014). The functional consequences of variation in transcription factor binding. PLoS Genet. 10, e1004226. [PubMed: 24603674]

Depew MJ, Liu JK, Long JE, Presley R, Meneses JJ, Pedersen RA, and Rubenstein JL (1999). Dlx5 regulates regional development of the branchial arches and sensory capsules. Development 126, 3831–3846. [PubMed: 10433912]

Dickel DE, Ypsilanti AR, Pla R, Zhu Y, Barozzi I, Mannion BJ, Khin YS, Fukuda-Yuzawa Y, Plajzer-Frick I, Pickle CS, et al. (2018). Ultracon-served enhancers are required for normal development. Cell 172, 491–499.e15. [PubMed: 29358049]

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. [PubMed: 23104886]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Eisenstat DD, Liu JK, Mione M, Zhong W, Yu G, Anderson SA, Ghattas I, Puelles L, and Rubenstein JL (1999). DLX-1, DLX-2, and DLX-5 expression define distinct stages of basal forebrain differentiation. J. Comp. Neurol 414, 217–237. [PubMed: 10516593]

Ernst J, and Kellis M (2017). Chromatin-state discovery and genome annotation with ChromHMM. Nat. Protoc 12, 2478–2492. [PubMed: 29120462]

Feledy JA, Morasso MI, Jang SI, and Sargent TD (1999). Transcriptional activation by the homeodomain protein distal-less 3. Nucleic Acids Res. 27, 764–770. [PubMed: 9889271]

Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C, Panning B, Ploegh HL, Bassik MC, et al. (2014). Genome-scale CRISPR-mediated control of gene repression and activation. Cell 159, 647–661. [PubMed: 25307932]

Gu Z (2015). rGREAT: client for GREAT analysis. https://github.com/jokergoo/rGREAT.

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, and Glass CK (2010). Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. Mol. Cell 38, 576–589. [PubMed: 20513432]

Hnisz D, Shrinivas K, Young RA, Chakraborty AK, and Sharp PA (2017). A phase separation model for transcriptional control. Cell 169, 13–23. [PubMed: 28340338]

Hoch RV, Lindtner S, Price JD, and Rubenstein JLR (2015). OTX2 transcription factor controls regional patterning within the medial ganglionic eminence and regional identity of the septum. Cell Rep. 12, 482–494. [PubMed: 26166575]

Jostes B, Walther C, and Gruss P (1990). The murine paired box gene, Pax7, is expressed specifically during the development of the nervous and muscular system. Mech. Dev 33, 27–37. [PubMed: 1982921]

Kothary R, Clapoff S, Darling S, Perry MD, Moran LA, and Rossant J (1989). Inducible expression of an hsp68-lacZ hybrid gene in transgenic mice. Development 105, 707–714. [PubMed: 2557196]

Krueger F (2015). Trim Galore!: a wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files (Babraham Institute).

Kuwajima T, Nishimura I, and Yoshikawa K (2006). Necdin promotes GABAergic neuron differentiation in cooperation with Dlx homeodomain proteins. J. Neurosci 26, 5383–5392. [PubMed: 16707790]

Li H, and Durbin R (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. [PubMed: 19451168]

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, and Durbin R; 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079. [PubMed: 19505943]

Liao Y, Smyth GK, and Shi W (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930. [PubMed: 24227677]

Lim L, Mi D, Llorca A, and Marín O (2018). Development and functional diversification of cortical interneurons. Neuron 100, 294–313. [PubMed: 30359598]

Liu JK, Ghattas I, Liu S, Chen S, and Rubenstein JL (1997). Dlx genes encode DNA-binding proteins that are expressed in an overlapping and sequential pattern during basal ganglia differentiation. Dev. Dyn 210, 498–512. [PubMed: 9415433]

Long JE, Garel S, Alvarez-Dolado M, Yoshikawa K, Osumi N, Alvarez-Buylla A, and Rubenstein JLR (2007). Dlx-dependent and -independent regulation of olfactory bulb interneuron differentiation. J. Neurosci 27, 3230–3243. [PubMed: 17376983]

Long JE, Cobos I, Potter GB, and Rubenstein JLR (2009a). Dlx1&2 and Mash1 transcription factors control MGE and CGE patterning and differentiation through parallel and overlapping pathways. Cereb. Cortex 19 (Suppl 1), i96–i106. [PubMed: 19386638]

Long JE, Swan C, Liang WS, Cobos I, Potter GB, and Rubenstein JLR (2009b). Dlx1&2 and Mash1 transcription factors control striatal patterning and differentiation through parallel and overlapping pathways. J. Comp. Neurol 512, 556–572. [PubMed: 19030180]

Magklara A, Yen A, Colquitt BM, Clowney EJ, Allen W, Markenscoff-Papadimitriou E, Evans ZA, Kheradpour P, Mountoufaris G, Carey C, et al. (2011). An epigenetic signature for monoallelic olfactory receptor expression. Cell 145, 555–570. [PubMed: 21529909]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Mayer C, Hafemeister C, Bandler RC, 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, 457–462. [PubMed: 29513653]

Mi D, Li Z, Lim L, Li M, Moissidis M, Yang Y, Gao T, Hu TX, Pratt T, Price DJ, et al. (2018). Early emergence of cortical interneuron diversity in the mouse embryo. Science 360, 81–85. [PubMed: 29472441]

Mowat DR, Wilson MJ, and Goossens M (2003). Mowat-Wilson syndrome. J. Med. Genet 40, 305–310. [PubMed: 12746390]

Niceta M, Stellacci E, Gripp KW, Zampino G, Kousi M, Anselmi M, Traversa A, Ciolfi A, Stabley D, Bruselles A, et al. (2015). Mutations impairing GSK3-mediated MAF phosphorylation cause cataract, deafness, intellectual disability, seizures, and a Down syndrome-like facies. Am. J. Hum. Genet 96, 816–825. [PubMed: 25865493]

Olivetti PR, and Noebels JL (2012). Interneuron, interrupted: molecular pathogenesis of ARX mutations and X-linked infantile spasms. Curr. Opin. Neurobiol 22, 859–865. [PubMed: 22565167]

Pai EL-L, Vogt D, Clemente-Perez A, McKinsey GL, Cho FS, Hu JS, Wimer M, Paul A, Fazel Darbandi S, Pla R, et al. (2019). Mafb and c-Maf have prenatal compensatory and postnatal antagonistic roles in cortical interneuron fate and function. Cell Rep. 26, 1157–1173.e5. [PubMed: 30699346]

Panganiban G, and Rubenstein JLR (2002). Developmental functions of the Distal-less/Dlx homeobox genes. Development 129, 4371–4386. [PubMed: 12223397]

Pattabiraman K, Golonzhka O, Lindtner S, Nord AS, Taher L, Hoch R, Silberberg SN, Zhang D, Chen B, Zeng H, et al. (2014). Transcriptional regulation of enhancers active in protodomains of the developing cerebral cortex. Neuron 82, 989–1003. [PubMed: 24814534]

Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, Shoukry M, Minovitsky S, Dubchak I, Holt A, Lewis KD, et al. (2006). In vivo enhancer analysis of human conserved non-coding sequences. Nature 444, 499–502. [PubMed: 17086198]

Petryniak MA, Potter GB, Rowitch DH, and Rubenstein JLR (2007). Dlx1 and Dlx2 control neuronal versus oligodendroglial cell fate acquisition in the developing forebrain. Neuron 55, 417–433. [PubMed: 17678855]

Phipson B, Lee S, Majewski IJ, Alexander WS, and Smyth GK (2016). Robust hyperparameter estimation protects against hypervariable genes and improves power to detect differential expression. Ann. Appl. Stat 10, 946–963. [PubMed: 28367255]

Pla R, Stanco A, Howard MA, Rubin AN, Vogt D, Mortimer N, Cobos I, Potter GB, Lindtner S, Price JD, et al. (2018). Dlx1 and Dlx2 promote interneuron GABA synthesis, synaptogenesis, and dendritogenesis. Cereb. Cortex 28, 3797–3815. [PubMed: 29028947]

Püschel AW, and Betz H (1995). Neurexins are differentially expressed in the embryonic nervous system of mice. J. Neurosci 15, 2849–2856. [PubMed: 7722633]

Qiu M, Bulfone A, Martinez S, Meneses JJ, Shimamura K, Pedersen RA, and Rubenstein JL (1995). Null mutation of Dlx-2 results in abnormal morphogenesis of proximal first and second branchial arch derivatives and abnormal differentiation in the forebrain. Genes Dev. 9, 2523–2538. [PubMed: 7590232]

Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, and Manke T (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 44 (W1), W160–W165. [PubMed: 27079975]

Rétaux S, Rogard M, Bach I, Failli V, and Besson M-J (1999). Lhx9: a novel LIM-homeodomain gene expressed in the developing forebrain. J. Neurosci 19, 783–793. [PubMed: 9880598]

Robinson MD, McCarthy DJ, and Smyth GK (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140. [PubMed: 19910308]

Rosenbluh J, Xu H, Harrington W, Gill S, Wang X, Vazquez F, Root DE, Tsherniak A, and Hahn WC (2017). Complementary information derived from CRISPR Cas9 mediated gene deletion and suppression. Nat. Commun 8, 15403. [PubMed: 28534478]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Sandberg M, Flandin P, Silberberg S, Su-Feher L, Price JD, Hu JS, Kim C, Visel A, Nord AS, and Rubenstein JLR (2016). Transcriptional networks controlled by NKX2–1 in the development of forebrain GABAergic neurons. Neuron 91, 1260–1275. [PubMed: 27657450]

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, and Ideker T (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498–2504. [PubMed: 14597658]

Shen L, Shao N, Liu X, and Nestler E (2014). ngs. plot: quick mining and visualization of next-generation sequencing data by integrating genomic databases. BMC Genomics 15, 284. [PubMed: 24735413]

Silberberg SN, Taher L, Lindtner S, Sandberg M, Nord AS, Vogt D, Mckinsey GL, Hoch R, Pattabiraman K, Zhang D, et al. (2016). Subpallial enhancer transgenic lines: a data and tool resource to study transcriptional regulation of GABAergic cell fate. Neuron 92, 59–74. [PubMed: 27710791]

Tasic B, Hippenmeyer S, Wang C, Gamboa M, Zong H, Chen-Tsai Y, and Luo L (2011). Site-specific integrase-mediated transgenesis in mice via pronuclear injection. Proc. Natl. Acad. Sci. USA 108, 7902–7907. [PubMed: 21464299]

Thakore PI, D'Ippolito AM, Song L, Safi A, Shivakumar NK, Kabadi AM, Reddy TE, Crawford GE, and Gersbach CA (2015). Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. Nat. Methods 12, 1143–1149. [PubMed: 26501517]

Toresson H, and Campbell K (2001). A role for Gsh1 in the developing striatum and olfactory bulb of Gsh2 mutant mice. Development 128, 4769–4780. [PubMed: 11731457]

Visel A, Taher L, Girgis H, May D, Golonzhka O, Hoch RV, McKinsey GL, Pattabiraman K, Silberberg SN, Blow MJ, et al. (2013). A high-resolution enhancer atlas of the developing telencephalon. Cell 152, 895–908. [PubMed: 23375746]

Vokes SA, Ji H, McCuine S, Tenzen T, Giles S, Zhong S, Longabaugh WJR, Davidson EH, Wong WH, and McMahon AP (2007). Genomic characterization of Gli-activator targets in sonic hedgehog-mediated neural patterning. Development 134, 1977–1989. [PubMed: 17442700]

Wang Y, Dye CA, Sohal V, Long JE, Estrada RC, Roztocil T, Lufkin T, Deisseroth K, Baraban SC, and Rubenstein JLR (2010). Dlx5 and Dlx6 regulate the development of parvalbumin-expressing cortical interneurons. J. Neurosci 30, 5334–5345. [PubMed: 20392955]

Wang B, Lufkin T, and Rubenstein JLR (2011). Dlx6 regulates molecular properties of the striatum and central nucleus of the amygdala. J. Comp. Neurol 519, 2320–2334. [PubMed: 21452241]

Wang B, Long JE, Flandin P, Pla R, Waclaw RR, Campbell K, and Rubenstein JLR (2013). Loss of Gsx1 and Gsx2 function rescues distinct phenotypes in Dlx1/2 mutants. J. Comp. Neurol 521, 1561–1584. [PubMed: 23042297]

Xu Z, Liang Q, Song X, Zhang Z, Lindtner S, Li Z, Wen Y, Liu G, Guo T, Qi D, et al. (2018). SP8 and SP9 coordinately promote D2-type medium spiny neuron production by activating Six3 expression. Development 145, dev.165456.

Yun K, Potter S, and Rubenstein JL (2001). Gsh2 and Pax6 play complementary roles in dorsoventral patterning of the mammalian telencephalon. Development 128, 193–205. [PubMed: 11124115]

Yun K, Fischman S, Johnson J, Hrabe de Angelis M, Weinmaster G, and Rubenstein JLR (2002). Modulation of the notch signaling by Mash1 and Dlx1/2 regulates sequential specification and differentiation of progenitor cell types in the subcortical telencephalon. Development 129, 5029–5040. [PubMed: 12397111]

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, and Liu XS (2008). Model-based analysis of ChIP-seq (MACS). Genome Biol. 9, R137. [PubMed: 18798982]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Highlights

- DLX proteins modulate expression of GABAergic neuronal differentiation genes
- DLXs drive a complex regulatory network by transcriptional activation and inhibition
- Genomic and epigenomic context and sequence predict DLX effect on gene expression
- DLX regulatory wiring may reflect general mechanisms implicated in neurodevelopment

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

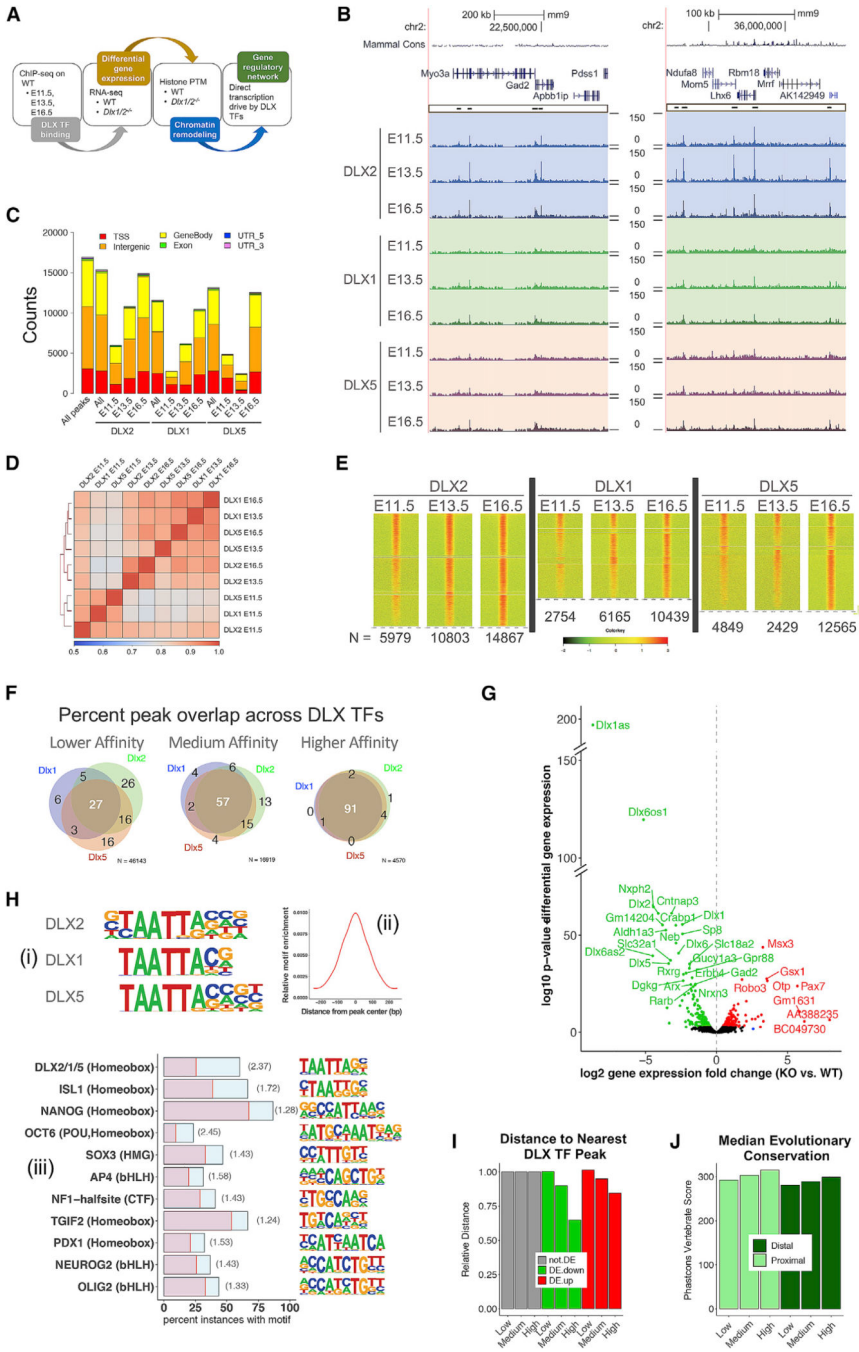


Figure 1. DLX2, DLX1, and DLX5 Genomic Binding in Embryonic GE

(A) Schematic of functional genomic dissection of DLX.

(B) DLX ChIP-seq coverage at *Gad2* and *Lhx6* loci; merged peak dataset represented inside a golden box on top of first DLX track.

(C) DLX peak counts by genomic feature at E11.5, E13.5, and E16.5.

(D) Heatmap showing pairwise Pearson correlation for genome-wide coverage values for DLX ChIP-seq.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(E) Normalized coverage of ChIP-seq peaks. Each row represents a DLX binding region ± 10 kb. Numbers under the heatmap columns denote number of peaks called for each DLX/time point.

(F) Venn diagrams showing increasing percent of peaks shared across DLXs as peak stringency increases.

(G) Volcano plot showing E13.5 *Dlx1/2*$^{-/-}$ versus WT GE differential gene expression fold change and statistical significance.

(H) Shared primary binding DNA motif across DLXs (i) centered within ChIP-seq peaks (ii). Motifs that were strongly enriched within DLX peaks (iii); blue bars show motif frequency in DLX peaks and red bars in GC-matched background sequences (enrichment in parentheses).

(I) Distance from TSS to nearest DLX peak for DE genes by peak stringency.

(J) Median Phastcons scores for vertebrate-conserved elements at promoter and distal DLX peaks by stringency.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

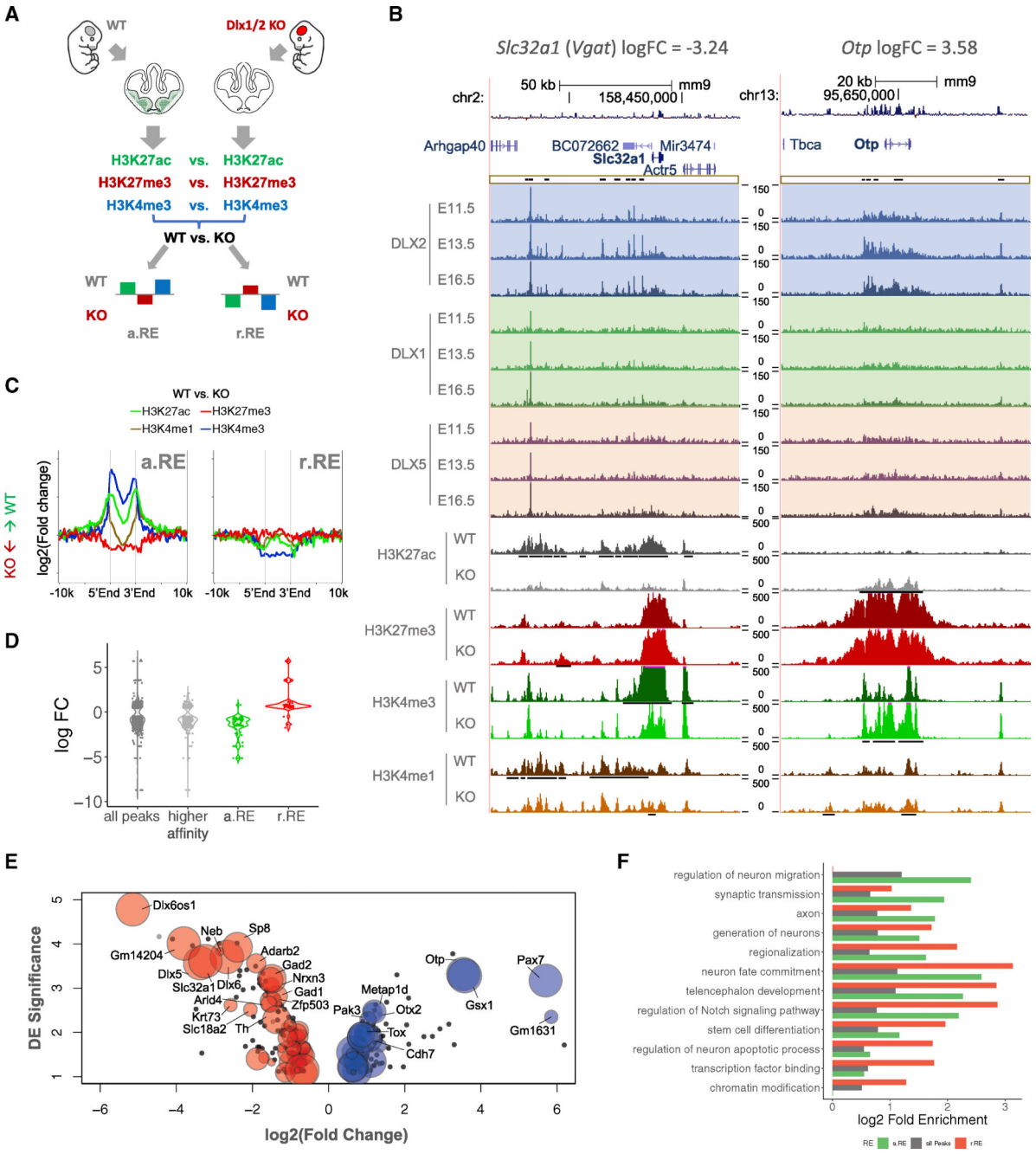


Figure 2. Chromatin State Is Dependent on DLX Binding for Key Regulatory Targets

(A) Schematic of epigenomic comparison of WT and *Dlx1/2*$^{-/-}$ E13.5 GE.

(B) DLX and histone ChIP-seq coverage for *Slc32a1* (*Vgat,* downregulated) and *Otp* (upregulated) loci; merged peak dataset represented inside a golden box on top of first DLX track, and differential histone PTMs shown as black bars under the peak where there was statistically significant difference.

(C) Average change in histone PTM signal in *Dlx1/2*$^{-/-}$ GE at DLX-bound loci for regions featuring loss of activating or gain of repressive marks (a.RE) or gain of activating or loss of repressive marks (r.RE).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(D) Log_2 fold change of nearest gene expression for all DLX peaks, higher-affinity peaks, and a.RE and r.RE loci.
(E) Modified volcano plot showing intersection of all differential gene expression and a.RE and r.RE loci. Colored circles represent a.RE (red) or r.RE (blue) within 100 kb of gene TSS. Circle size shows magnitude of histone PTM change in *Dlx1/2*$^{-/-}$ GE.
(F) Differential enrichment of functional annotation terms comparing all DLX TF peaks and a.RE and r.RE loci.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

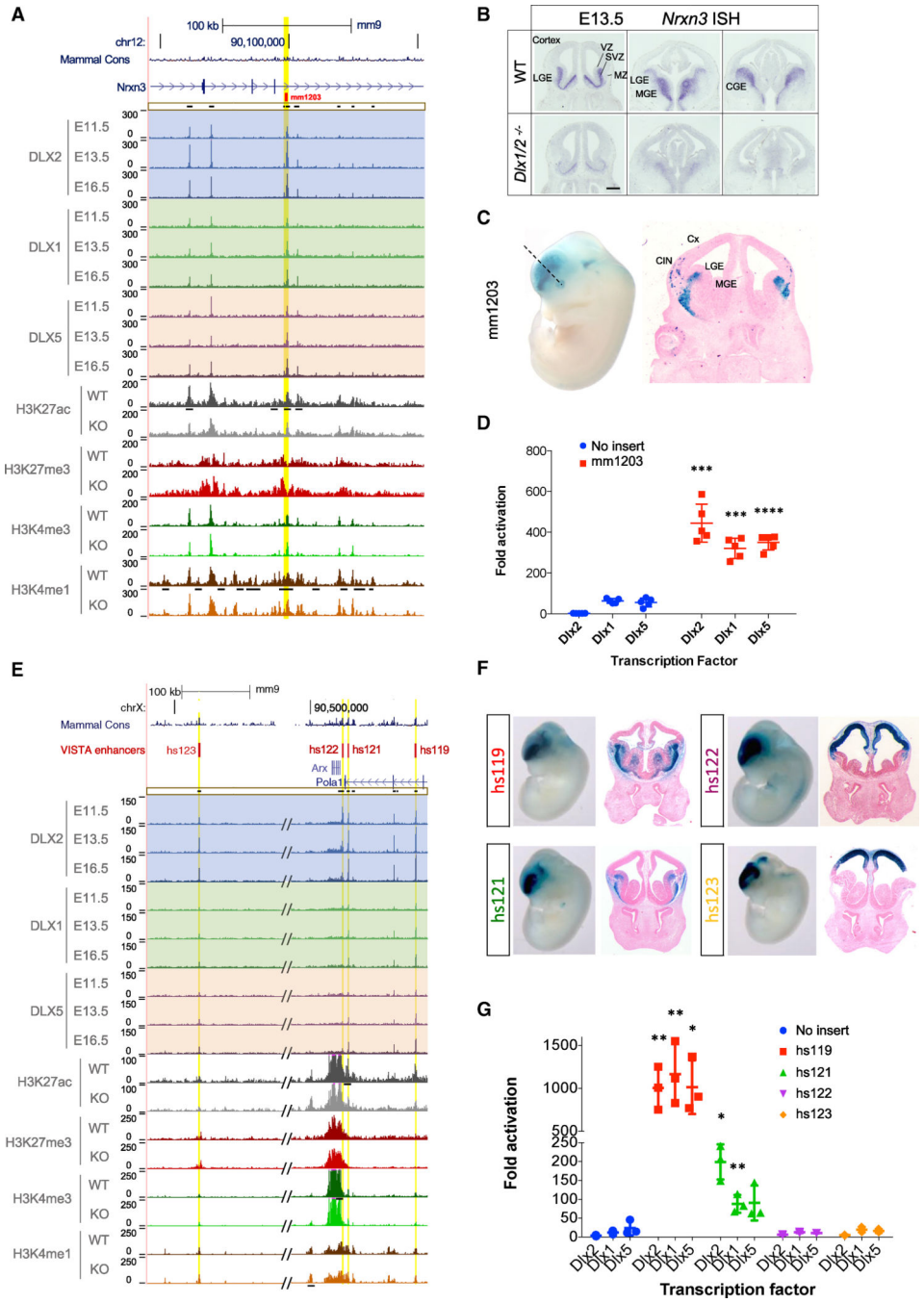


Figure 3. *Nrxn3* and *Arx* Are Regulated by DLX via RE Interactions

(A) ChIP-seq data for *Nrxn3* locus; called DLX peaks and histone differential PTM as described in Figure 2B.

(B) ISH analysis of *Nrxn3* expression in WT (upper panel) and *Dlx1/2*$^{-/-}$ (lower panel) forebrain at E13.5. (MGE, medial GE; LGE, lateral GE; CGE, caudal GE; SVZ, subventricular zone; VZ, ventricular zone; MZ, mantle zone). Images are representatives of 3 embryos. Scale bar represents 500 μm.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(C) *mm1203* sequence drives *LacZ* expression in E12.5 forebrain in a transgenic mouse enhancer assay. Dashed line indicates plane of section.
(D) *Dlx* transfection luciferase transcription assays in P19 cells showing activating effect of DLX on *mm1203.*
(E) ChIP-seq data for *Arx* locus; called DLX peaks and histone differential PTM as described in Figure 2B.
(F) Activity patterns in mouse transgenic assays for characterized *Arx* GE (*hs119, hs121*) and pallial (*hs122, hs123*) REs.
(G) *Dlx* transfection luciferase transcription assays in P19 cells showing specific activating effect of DLX on *Arx* GE REs. Luciferase data are represented as mean ± SEM (n = 5 for *Nrxn3* and n = 3 for *Arx*). Unpaired t test was used for the statistical analysis between the presence of enhancer in the reporter or no enhancer (*p < 0.05, **p <0.01, ***p <0.001, and ****p < 0.0001).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

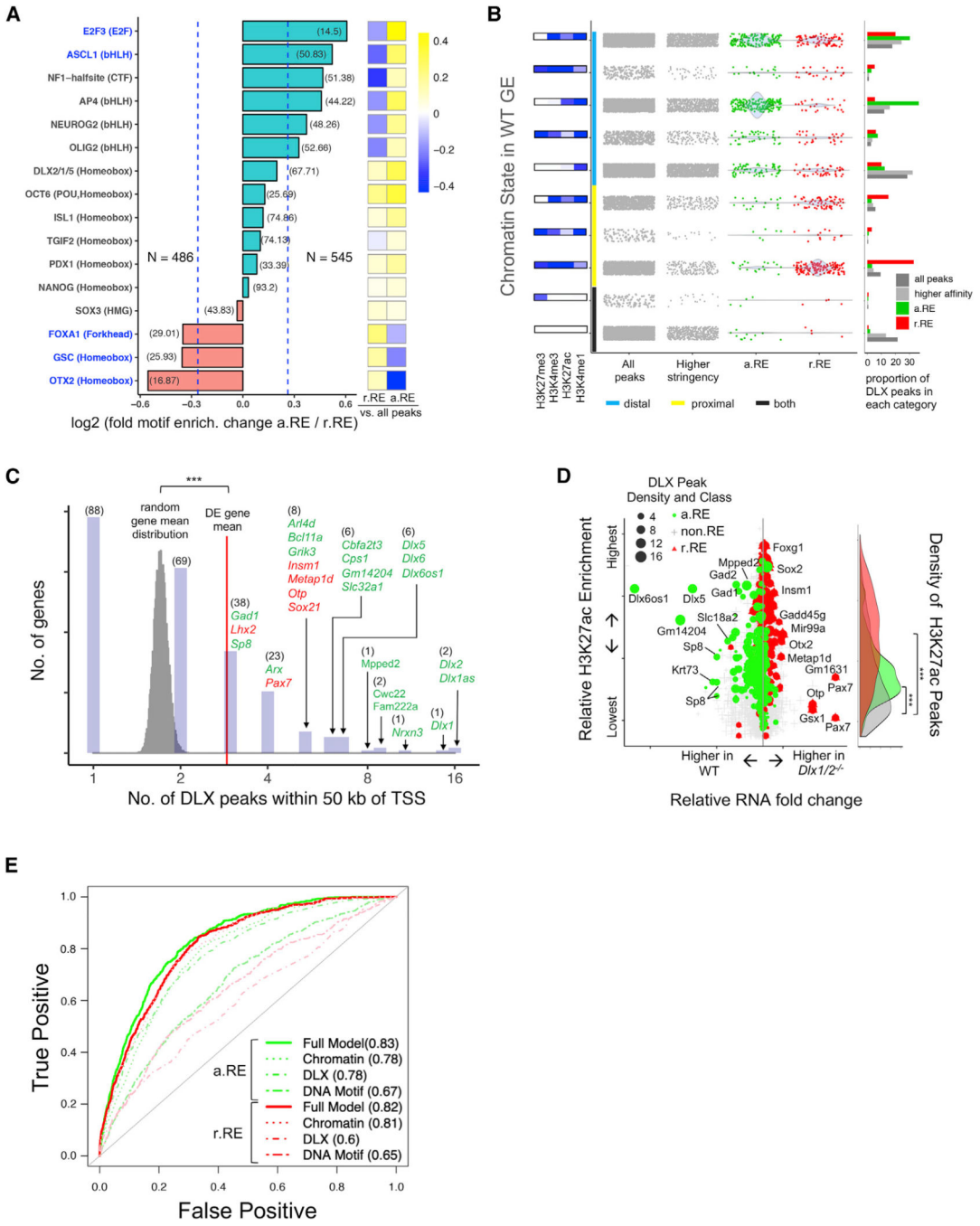


Figure 4. Genome Sequence and Context Define a.RE and r.RE Loci

(A) Differential motif enrichment in a.RE and r.RE. loci. Motifs specific to a.RE or r.RE labeled in blue. Horizontal bar plot on left shows $\log_2$ fold enrichment difference between a.RE and r.RE with dotted blue lines shown for 1.2-fold enrichment difference. Heatmap on right shows a.RE and r.RE enrichment versus all DLX peaks.

(B) Distribution of all DLX peaks, higher-affinity peaks, and a.RE and r.RE loci by chromatin state. ChromHMM states defined by relative histone PTM enrichment as shown on left. Each dot in center plot represents one peak with density represented by shaded area.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Distal and proximal (promoter) region assignment shown as colored bars along they axis, with blue and yellow indicating distal and proximal regions, respectively; no chromatin state signal and repressed regions marked as both (black). Proportion of total peaks shown in bar plot at right.

(C) Distribution of the number of DLX peaks within 50 kb of TSS for all versus DE genes, with gene counts by DLX peak number histogram and representative example genes shown.

(D) Relationship between $\log_2$ FC (fold change of gene expression *Dlx1/2*$^{-/-}$ versus WT) of the nearest TSS (x-axis) and histone H3K27ac signal in WT E13.5 GE (y-axis) and for all DLX peaks, a.RE, and r.RE. Symbol size shows density of DLX TF peaks within 50 kb. Distribution of H3K27ac enrichment for peak classes shown at right.

(E) Receiver operating characteristic (ROC) curve for logistic regression models predicting a.RE or r.RE status. Predictors include DNA motif presence, chromatin context (histone PTM, distal/proximal, ChromHMM state) and DLX (affinity and local peak density), and a full model with all features included.

***$p < 2.2 \times 10^{-16}$ in (C) and (D).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

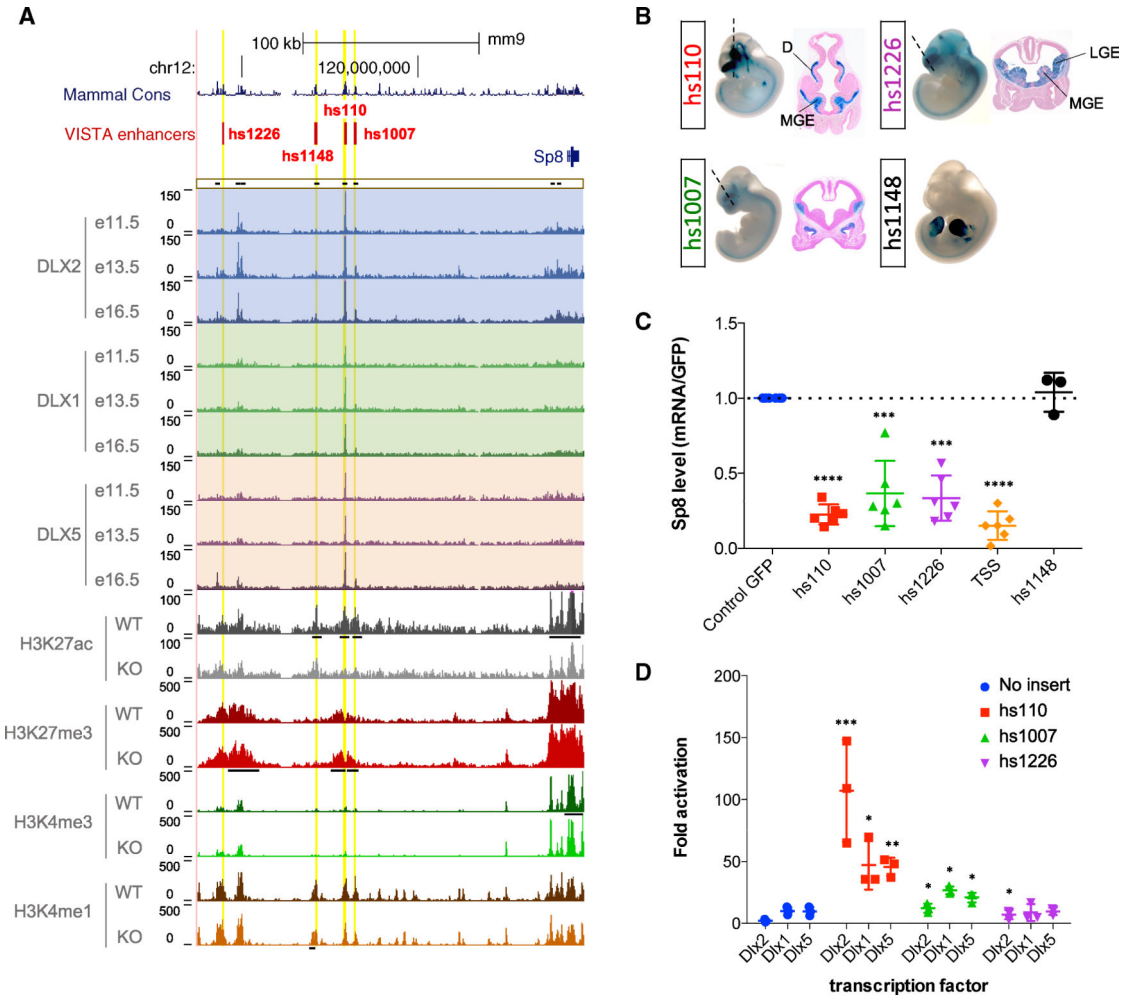


Figure 5. *Sp8* Is Regulated by DLX Binding to Promoter-Interacting Distal Elements

(A) ChIP-seq coverage for *Sp8* locus; called DLX peaks and histone differential PTM as described in Figure 2B.

(B) Transgenic RE assay results showing activity of three GE-active (*hs110, hs1226,* and *hs1007*) and limb-active (*hs1148*) REs.

(C) Downregulation of *Sp8* mRNA caused by delivery of dCas9-KRAB and gRNA targeting putative *Sp8* REs. y-axis shows the amount of *Sp8* mRNA relative to control GFP gRNA (blue points) measured in RT-qPCR (n = 3 for gRNA-*hs1148*; n = 6 for the rest).

(D) *Dlx* transfection luciferase transcription assay in P19 cells showing the effect of DLX on the *Sp8* GE REs; data represented as mean ± SEM (n = 3). Unpaired t test was used for the statistical analysis between the presence of enhancer in the reporter or no enhancer (*p < 0.05; **p < 0.01, ***p < 0.001, ****p < 0.0001).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

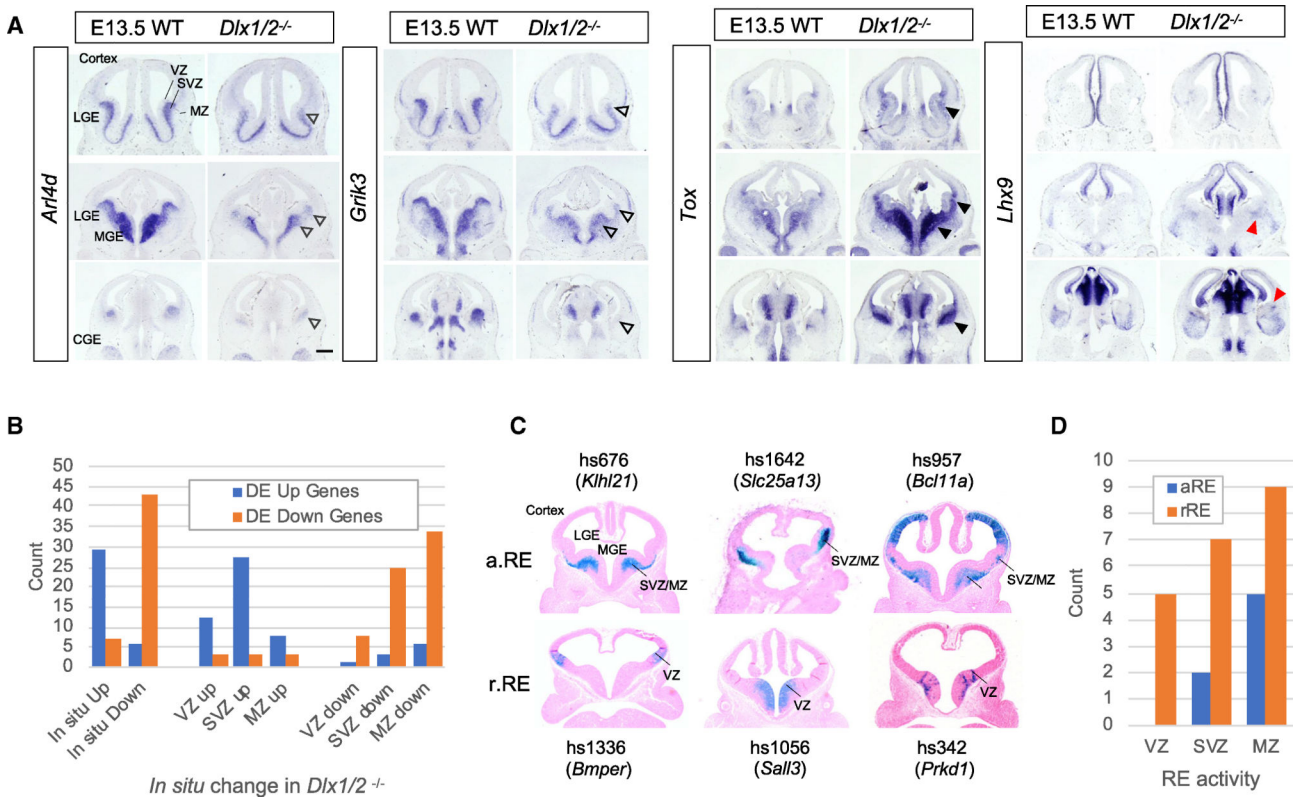


Figure 6. DLX TFs Regulate Spatial Patterns of Gene Expression

(A) ISH analyses using probes detecting *Arl4d* and *Grik3,* and *Tox* and *Lhx9* in WT and *Dlx1/2*$^{-/-}$ forebrain at E13.5 (MGE, medial GE; LGE, lateral GE; CGE, caudal GE; SVZ, subventricular zone; VZ, ventricular zone; MZ, mantle zone; open arrows, downregulation; black arrows, upregulations; red arrows, ectopic expression). Images are representatives of 2 embryos for each probe. Scale bar represents 500 μm.

(B) Summary of spatial changes in DE gene expression from ISH analysis comparing WT and *Dlx1/2*$^{-/-}$ forebrain.

(C) Transgenic RE assay activity patterns for representative a.RE and r.RE sequences.

(D) Summary of RE activity patterns within GE subregions (VZ, ventricular zone; SVZ, subventricular zone; MZ, mantle zone).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

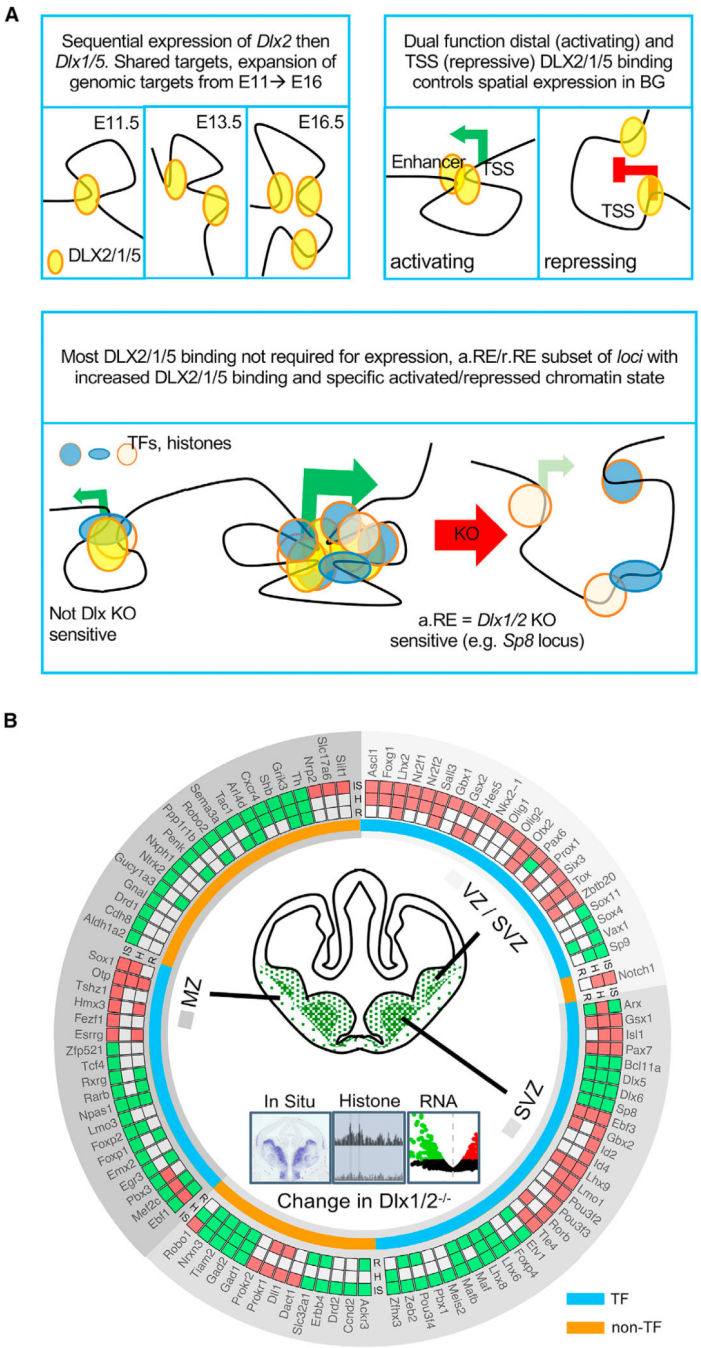


Figure 7. Regulatory Model and Gene Regulatory Networks Orchestrated by DLX

(A) Schematic model of DLX genomic function in developing GE.

(B) Curated gene regulatory network (GRN) for DLX-regulated transcription factors and lineage specification factors. The GRN is organized radially with regard to the laminar activity of genes (VZ/SVZ, SVZ, and MZ, shown in three different shades of gray). The effects of the *Dlx1/2*$^{-/-}$ mutation are indicated in three nested circles. The outer circle reports RNA changes measured by ISH (IS) (this study; Long et al., 2009a, 2009b). The middle circle reports histone PTM signal (H) changes. The inner circle reports RNA-seq (R)

changes. Red and green represent repressive and activating roles for DLX TFs on each assay, respectively. For histone changes, REs assigned to nearest TSS.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
DLX2 (MBP-aa1–154)	This paper	N/A; Available from the authors
DLX1 (His-aa1–121)	This paper	N/A; Available from the authors
DLX5 (His-188–289aa)	This paper	N/A; Available from the authors
H3K4me1	Abcam	Cat# ab8895; RRID:AB_306847
H3K4me3	Abcam	Cat# ab8580; RRID:AB_306649
H3K27me3	Active Motif	Cat# 39155; RRID:AB_2561020
H3K27ac	Abcam	Cat# ab4729:RRID:AB_2118291
Anti-Digoxigenin-AP	Sigma-Aldrich	Cat# 11093274910; RRID:AB_2734716
Goat anti-Rabbit IgG (H+L), Alexa Fluor 546	Thermo-Fisher	Cat# A-11035; RRID:AB_2534093
Normal Rabbit IgG	Santa Cruz	Cat# sc-2027; RRID:AB_737197
Chemicals, Peptides and Recombinant Proteins		
DIG RNA labeling mix	Sigma-Aldrich	Cat# 11277073910
T7 RNA polymerase	Sigma-Aldrich	Cat# 10881767001
DNaseI	Sigma-Aldrich	Cat# 10104159001
Sheep Serum	Sigma-Aldrich	Cat# S2263
Blocking reagent	Sigma-Aldrich	Cat# 11096176001
BM purple	Sigma-Aldrich	Cat# 11442074001
37% Formaldehyde	Ted Pella	Cat# 18508
Dynabeads Protein G	Thermo-Fisher	Cat# 10003D
Dynabeads Protein A	Thermo-Fisher	Cat# 10001D
RNase, DNase free	Sigma-Aldrich	Cat# 11119915001
Proteinase K	Sigma-Aldrich	Cat# 3115879001
PerfeCTa SYBR Green FastMix ROX	Quanta	Cat# 95073–012
MNase	Sigma-Aldrich	Cat# N3755–50UN
Thermo Scientific Shandon Aqua-Mount Slide Mounting Media	Fisher Scientific	Cat# 14–390–5
X-tremeGENE HP DNA transfection reagent	Sigma-Aldrich	Cat# 6366236001
Complete EDTA-free Protease inhibitor	Sigma-Aldrich	Cat# 11873580001
Critical Commercial Assays		
RNA Clean and Concentrator	Zymo Research	Cat# R1015
Ovation Ultralow DR Multiplex System	Nugen	Cat# 0344–32
2% Agarose Gel Cassette Blue Pippin	Sage Science	Cat# BDF2010
High Sensitivity DNA Reagents	Agilent Technologies	Cat# 5067–4626
TruSeq Stranded Total RNA Library Prep Kit with Ribo-Zero Mouse	Illumina	Cat# RS-122–2202
KAPA Library Quantification Kit	Roche	Cat# 07960140001

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

REAGENT or RESOURCE	SOURCE	IDENTIFIER
RNeasy Mini Kit	QIAGEN	Cat# 74104
NE-PER kit	Thermo Fisher	Cat# 78833
Dual Luciferase Reporter Assay System	Promega	Cat# E1980
Deposited Data		
RNA-seq and ChIP-seq data	US National Center for Biotechnology Information	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE124936
Recombinant DNA		
Hsp68-lacZ vector	PMID 2557196	n/a (available from DED, AV and LAP)
psPax2	Addgene	Cat# 12260
pmD2G	Addgene	Cat# 12259
Experimental Models: Organisms/ Strains		
Transgenic enhancer assay: FVB	Charles River	https://www.criver.com/
ChIP-seq: CD1	Charles River	https://www.criver.com/
dCas9-KRAB	V.G., M. Lebedinskaya, R. Wagner, R. Jaafar, M.T.M., unpublished data	n/a
Software and Algorithms		
STAR	Dobin et al., 2013	https://github.com/alexdobin/STAR
FASTQC	Andrews, 2010	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
featureCounts	Liao et al., 2014	http://subread.sourceforge.net
edgeR	Phipson et al., 2016	https://www.bioconductor.org/packages/release/bioc/html/edgeR.html
limma	Robinson et al., 2010	https://www.bioconductor.org/packages/release/bioc/html/limma.html
Trim Galore	Babraham Bioinformatics	https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/
BWA	Li and Durbin, 2009	http://bio-bwa.sourceforge.net/bwa.shtml
Samtools	Li et al., 2009	http://samtools.sourceforge.net
DeepTools	Ramirez et al., 2016	https://deeptools.readthedocs.io/en/develop/
MACS2	Zhang et al., 2008	https://github.com/taoliu/MACS/
ngsplot	Shen et al., 2014	https://github.com/shenlab-sinai/ngsplot
ChromHMM	Ernst and Kellis, 2017	http://compbio.mit.edu/ChromHMM/
HOMER	Heinz et al., 2010	http://homer.ucsd.edu/homer/
Cytoscape	Shannon et al., 2003	https://cytoscape.org
rGREAT	Gu, 2015	https://www.bioconductor.org/packages/release/bioc/html/rGREAT.html