# Chromatin landscapes of the *Dlx1/2* and *Dlx5/6* bigene clusters in the developing mouse forebrain

Simon Monis

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies in partial fulfillment of the requirement for the Master of Science degree in Biology

Department of Biology

Faculty of Science

University of Ottawa

**Abstract**

The *Distal-less (Dlx)* homeobox genes of mammals are expressed in many tissues of the developing organism including the limbs, craniofacial skeleton and the forebrain. In the forebrain, *Dlx1, Dlx2*, *Dlx5* and *Dlx6* play a critical role in driving tangential migration of GABAergic progenitors from the ventral telencephalon to their final destinations, notably the neocortex and the striatum. These *Dlx* genes are organised into convergently transcribed clusters with short intergenic regions that contain notable cis regulators elements (CREs) that drive *Dlx* expression in unique subdomains of the developing ventral telencephalon. Previous studies have characterised *Dlx* regulation including but not limited to the direct activation of these CREs by effector proteins. However, to date very little work has been done to examine how the forebrain *Dlx* genes may be regulated at the level of the chromatin. To explore this, I used *in silico* and *in vivo* methods to examine some key histone modifications of the *Dlx1/2* and *Dlx5/6* bigene clusters in the developing forebrain; namely H3K27Ac, H3K4me3, H3K4me1 and H3K27me3. I found that within the *Dlx* expressing ganglionic eminences (GE), at midgestation, the *Dlx* loci are marked by bivalent chromatin which is enriched in both permissive H3K4me3 and repressive H3K27me3 marks. By performing ChIP-qPCR on the GE tissue of embryonic mice with targeted deletions of enhancer CREs, I found that these CREs play unique roles in shaping the chromatin. Removal of one of these CREs has widespread effects on the chromatin at both loci. Since these changes in chromatin signatures do not accompany significant changes in expression of histone modifying genes, we believe these CREs play yet-to-be determined roles in recruiting the modifying proteins to the loci, thereby establishing bivalent chromatin to fine-tune *Dlx* expression.

**Résumé**

Les gènes *Dlx* des mammifères sont exprimés dans plusieurs tissus de l'organisme au cours du développement, incluant les membres, le squelette craniofacial, et le cerveau antérieur. Dans le cerveau antérieur, les gènes *Dlx1*, *Dlx2*, *Dlx5* et *Dlx6* jouent un rôle crucial dans le contrôle de la migration tangentielle des progéniteurs GABAergiques, du télencéphale ventral vers leurs destinations finales, notamment dans le néocortex et le striatum. Les gènes *Dlx* sont organisés en complexes qui sont transcrits de manière convergente et qui incluent une courte région intergénique. Celle-ci contient des éléments de régulation agissant en *cis* (CREs) qui contribuent à l'expression des *Dlx* dans des domaines uniques du télencéphale ventral au cours du développement. Des études antérieures ont montré que le contrôle de l'expression des *Dlx* dépendait, entre autres, de leur activation directe par des protéines régulatrices agissant en *trans*. Jusqu'à ce jour, peu d'attention a été apportée au rôle de la chromatine dans le contrôle de l'expression des *Dlx* dans le cerveau antérieur. Afin d'explorer ces mécanismes, j'ai utilisé des approches *in silico* et *in vivo* pour examiner le paysage chromatinique des complexes *Dlx1/2* et *Dlx5/6* dans le cerveau antérieur de l'embryon. J'ai observé que dans les éminences ganglionnaires (EG), à la mi-gestation, les loci *Dlx* sont marqués par de la chromatine bivalente qui est enrichie en H3K4me3 permissive et en H3K27me3 répressive. En utilisant la méthode ChIP-PCR sur du matériel obtenu d'EGs d'embryons de souris mutantes, dans lesquelles des CREs ont été retirés, j'ai observé que les CREs jouent un rôle unique dans l'organisation de la chromatine aux deux loci. Puisque ces changements de signatures chromatiniques ne sont pas accompagnés d'effets significatifs sur l'expression des protéines responsables des modifications des histones, je considère que les CREs jouent des rôles, encore

à déterminer, dans le recrutement des protéines modificatrices responsables de l'établissement d'une chromatine bivalente qui confère un contrôle fin de l'expression des *Dlx*.

**Acknowledgments**

This thesis was a labour of love. Doing science daily is a privilege and an honour, one which would not be possible without a great deal of time, patience and no small measure of support. I would like to take the time; however brief, to mention some of the people that made this work possible.

Firstly, I would like to thank my boss Dr. Marc Ekker. Marc is nothing short of a first-rate scientist and mentor. He has continuously been a source of wisdom and inspiration for me. I fell in love with not just his research program, but the concept of basic science research itself under his tutelage. Since research seems to be the direction my life is heading, it cannot be understated the role he has played in what I hope will be a successful carer in science. Thank you for taking a chance on me Marc.

Next, I would like to thank all the members of the Ekker lab. It is a great joy to come to the lab everyday knowing I will be surrounded by such passionate scientists, but better yet, genuinely warm human beings. They made each day more fun. They made the grind a little less excruciating. I would like to name a few of you who I now consider to be good friends: Michael Kalyn, Khang Hua, and Dũng Ngö. I would like to thank a visiting researcher in our lab, David Wong who spent a year with us. In that time, we spent a great deal of time together, including our fair share of uncannily late nights. We pushed ourselves to work hard and enjoy the process. I would also like to thank a former PhD in our lab Siavash Fazel Darbandi who taught me a lot about the ChIP assay and made me feel confident in my work. I would lastly like to give a special shout-out to our technician Gary Hatch. Despite being separated by more than 20

years I consider Gary to be one of the easiest individuals to talk to. We shared many conversations about our work, but also our lives outside of the lab. Thank each one of you for always listening to me ramble on about whatever thought I couldn't get out of my head that day.

Of course, I must thank all the people who make the work possible. This includes all the wonderful staff at the University of Ottawa - from building managers, to the science store employees - each one of them plays a roll in making this possible. I also would like to thank the members of ACVS who work and take care of the animals. They play a vital role in making the research process go smoothly. Thank you to Vishal Saxena in the fish room and Josh Lavigne in the mouse room for being my go-to animal experts. Thanks of course to CIHR for funding this work, and the work of thousands of research scientists across this country.

Finally, but most importantly to me, I would like to thank my family and friends for their support. I am pretty sure I won the lottery on this one. My family has constantly encouraged me to explore any avenue or interest I may have. Simply put, they wanted me to be happy and to try my best. The advantages of a support system like this are not lost on me. Thank you to my wonderful parents, my loving girlfriend, my good friends here in Ottawa and around the world, and my siblings. They all know the role they played in getting me here, and I thank them in advance for the role they will no doubt play in getting me to wherever I am going.

**Table of contents**

**List of figures**



**List of tables**

**Abbreviations and nomenclature**

3C – chromosome conformation capture

AP – action potential

AP axis – anterior-posterior axis

APE – anterior entopeduncular area

ASD – autistic spectrum disorder

*Ash1l – absent small and homeotic disk protein 1*

BAC – bacterial artificial chromosome

bp – base pairs

Brg1 – Brahma related gene-1

CGE – caudal ganglionic eminence

*Chd8 – chromodomain helicase DNA binding protein 8*

ChIP – chromatin immunoprecipitation

CNS – central nervous system

COMPASS - complex proteins associated with Set1

CREs – cis regulatory elements

*Dll – Distal-less*

*Dlx – Distal-less homeobox*

*Dlx1* P – D*lx1* promoter

*Dlx2* P – *Dlx2* promoter

*Dlx5* P – *Dlx5* promoter

*Dlx6* P – *Dlx6* promoter

DNA – deoxyribonucleic acid

E11.5 – embryonic development day 11.5 after conception

E13.5 – embryonic development day 13.5 after conception

E14.5 – embryonic development day 14.5 after conception

EDTA – ethylenediaminetetraacetic acid

EGTA - ethylene glycol-bis(β-aminoethyl ether)-*N*,*N*,*N*′,*N*′-tetraacetic acid

ENCODE – encyclopedia of DNA elements

ES cells – embryonic stem cells

*Evf2 – embryonic ventral forebrain-2*

GABA – γ-amino butyric acid

*Gad* – glutamic acid decarboxylase

GAD65 (GAD1) – GAD isoform 65 (GAD isoform 1)

GAD67 (GAD2) – GAD isoform 67 (GAD Isoform 2)

GAD3 – GAD isoform 3

GE – ganglionic eminence

Gtf2i – General transcription factor 2i

H2A – histone 2A

H2B – histone 2B

H3 – histone 3

H4 – histone 4

H3K4me1 – histone 3 lysine 4 trimethyl

H3K4me3 – histone 3 lysine 4 monomethyl

H3K9Ac – histone 3 lysine 9 acetyl

H3K27Ac – histone 3 lysine 27 acetyl

H3K27me3 – histone 3 lysine 27 trimethyl

H3K36me3 – histone 3 lysine 36 trimethyl

H3S10 – histone 3 serine 10

HAT – histone acetyl-transferases

HDAC – histone deacetylase

*Hdac1 – histone deactylase 1*

*Hdac2 -histone deacetylase 2*

*Hdac3 – histone deacetylase 3*

HMT – histone methyltransferases

*Hox – homeobox gene*

HR – homologous recombination

IP – immunoprecipitation

Kat2a – lysine acetyltransferase 2a

Kat2b – lysine acetyltransferase 2b

Kb – kilobases

*Kmt2c – lysine methyltransferase 2c*

LGE – lateral ganglionic eminence

LP – lateral plate

*Mash1 (Ascl1) – mammalian achaete-scute homolog-1 (achaete-scute family transcription factor 1)*

MGE – medial ganglionic eminence

MLL1 – mixed-lineage leukemia associated protein 1

MNase – micrococcal nuclease

mRNA – messenger ribonucleic acid

MZ - mantle zone

NRG1 – neuregulin 1

ns – not significant

NSC – neural stem cells

P0 – postnatal development day 0

P35 – postnatal development day 35

*Pax6 – paired box 6*

PBS – phosphate buffered saline

PgC – Polycomb group complex

PNS – peripheral nervous system

POE – preoptic area

qPCR – qualitative (real time) polymerase chain reaction

qRT-PCR – qualitative (real time) reverse transcription polymerase chain reaction

RNA - ribonucleic acid

rpm – rotations per minute

SD – standard deviation

SET - su(var)3-9, enhancer-of-zeste and trithorax

*Setd5 – set domain containing protein 5*

SEM – standard error of the mean

Shh – Sonic hedgehog

SNPS – single nucleotide polymorphisms

SVZ – subventricular zone

TCA – thalamocortical axons

TRX – Trithorax group proteins

URE2 – upstream response element 2

VZ – ventricular zone

WS – Williams syndrome

WT – wildtype

**Statement of contribution**

All the experiments performed in this monograph were conceived by Simon Monis, and Marc Ekker, except for the generation of the mutant mice which was completed by Luc Poitras. All the analysis and writing were done exclusively by Simon Monis.

# 1.Introduction

## 1.1 Chromatin and transcriptional regulation

### 1.1.1 Chromatin modification

The role of epigenetic alterations in controlling and fine-tuning gene expression has become increasingly evident in recent years. This is particularly true in developmental contexts when gene expression needs to be highly flexible and able to adapt quickly to both environmental and genetic cues (Reviewed in Bannister & Kouzarides, 2011). Epigenetic regulation includes: methylation of DNA sequences, histone modulation through many different chemical modifications including both acetylation and methylation, chromatin remodelling and interactions of trans or cis enhancer or promoter elements. These processes can be further complexified by interacting with each other, as well as other molecules including lncRNAs. Together these processes play a massive role in changing expression levels and expression domains of key developmental genes during embryogenesis. These processes can be said to play as significant a role in gene expression as conventional genetic-based systems like transcription factor activation and mRNA/protein processing. Indeed, some evidence suggests that a complete understanding of histone modifications be used to predict global transcriptomic profiles with a high degree of accuracy (Cheng et al., 2011).

The modulation of histones is a powerful epigenetic regulator of gene expression during development, in part because of its ability to modulate expression of many genes in parallel. This is a consequence of the way genes are organised in chromatin – wrapped around a nucleosome core particle every 147bp of DNA. The core particle contains a histone octamer

made of two copies of H2A, H2B, H3, and H4. While H2A and H2B can be targeted by modifications, these modifications are not as well understood or linked to direct transcriptional changes (H2A and H2B histone tail modifications are reviewed by Wyrick & Parra, 2009). The most studied and targeted histone modifications occur on H3. H4 tail modifications have also been closely linked to transcriptional control via modulating chromatin-protein interactions and chromatin stability among other things. (Shogren-Knaak et al., 2006; Oda et al., 2009). It is interesting to note that constitutively active promoters are often found in areas devoid of nucleosomes allowing transcription to occur totally uninhibited (nucleosome positioning and dynamics are reviewed in Rando & Ahmad, 2007)

It was traditionally thought that chromatin modification occurs through a complex system of modifying proteins and intermediates which are recruited to a given locus to put into affect a series of molecular changes to the chromatin. It is becoming increasingly evident however that several key developmental transcriptional regulators - TRX for example - in addition to affecting the transcription of their target genes, contain intrinsic histone-modifying domains (Katsani et al., 2001). In this way these transcription factors exhibit complex modalities and have effects that go beyond simple transcriptional activation of their immediate targets and can provide a two-pronged regulation system to both the genes they activate, as well as alter the transcription of genes they do not directly associate with. Furthermore, the lack of necessary protein intermediates may increase the kinetics of these epigenetic modifications relative to the traditional paradigm of recruiting modifying proteins and thus epigenetic regulation can occur in concert with transcriptional activation. It is important to note that while

many histone modifiers are basally expressed, their expression levels fluctuate widely during development and in different tissues.

### 1.1.2 Chromatin modifying proteins

Histone modulation is done by a series of proteins that are responsible for principally adding methyl groups or acetyl groups - frequently to lysine residues - through covalent linkage. Recent work has detailed the importance of other modifications including ubiquitination and phosphorylation - largely restricted to H3S10 - among others (histone ubiquination is reviewed in Cao & Yan, 2012; histone phosphorylation is reviewed in Sawicka & Seiser, 2012). All these modifying proteins are extraordinarily well conserved across kingdom animalia, with a high degree of overlap in functional outcome of these histone profiles - consequently a great deal of stability exists across evolution of the modified chromatin itself (Woo & Li, 2012).

These histone residue alterations do not occur on the nucleus of the histone fold, the areas where DNA is tightly associated with the nucleosome. Instead they occur on the N-terminal tails, unstructured sections of the core histones which do not closely associate to the DNA. These tails play a vital role in nucleosome-nucleosome interactions and in condensing the chromatin into heterochromatin for transcriptional depression (reviewed in Zheng & Hayes, 2003). In general, modifying proteins can be thought as tagging a histone tail. These altered histones can either be: 1) flagged for further processing, or 2) sterically altered by direct consequence of the modification (Bannister & Kouzarides, 2011). In the latter case, structural changes lead to an increase or decrease in access of the transcriptional machinery to the locus associated with the histone in question.

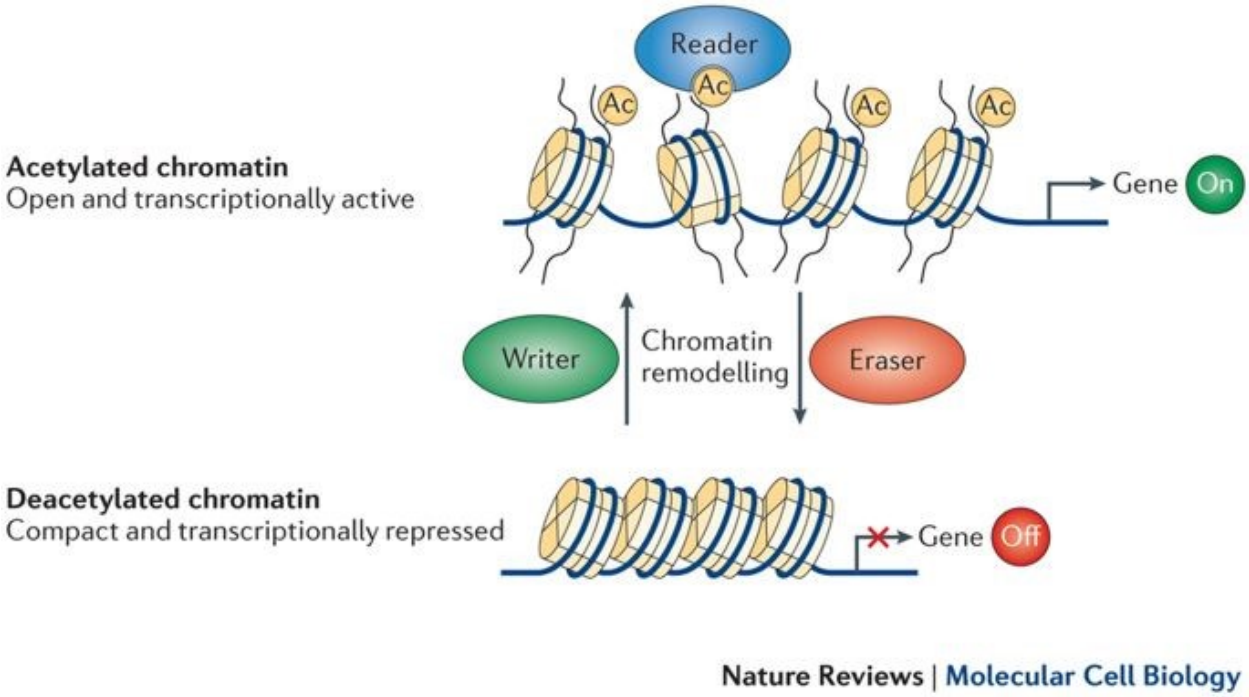
Methylation of histone residues is broadly achieved by a class of proteins called histone methyl transferases (HMTs). Many HMTs were first described as *Hox* regulators before their catalytic functions were first discovered (Milne et al., 2002). Histone methylation can be associated with both transcriptional activation and transcriptional repression of the surrounding loci. Histone methylators contain a common methylating domain called a SET domain which is highly conserved and contains a novel fold which both contains an active site and a chromatin stabilising site (Xiao, Wilson, & Gamblin, 2003).

The dual nature of methylated marks to be either activating or repressing is perhaps best illustrated in the context of the H3K4me3 modification versus H3K27me3 modification. The H3K4me3 mark is generally associated with transcriptional activation and is often highly enriched in actively transcribed genes; it is generally localised to transcription start sites (Barski et al., 2007). H3K4 methylation in general is completed by a complex of proteins called COMPASS proteins which are responsible for mono- di- and tri-methylations (Miller et al., 2001). This complex contains *Hox* regulators like Trithorax group proteins and SET proteins (COMPASS H3K4 methylases are reviewed in Shilatifard, 2012). These methylation marks are decoded by "reader proteins" which generally begin the process of nucleosome depletion allowing easy access of the transcriptional machinery to the DNA. H3K27me3 methylation, on the other hand, is associated with transcriptional silencing (Reviewed in Cheung & Lau, 2005 and by Wiles & Selker, 2017). Much work has been done on this front to illustrate the effects that this methylation plays in condensing chromatin around the H3K27me3 modulated nucleosome preventing access to the transcriptional machinery. This process is catalysed by *Hox* regulator Polycomb-group proteins that form large complexes (PRC1 and PRC2) including

the active proteins EZH1 and EZH2 which contain the SET domains (Polycomb protein activity and decoding is reviewed in Schuettengruber & Cavalli, 2009).

Unlike methylation, acetylation of histones is almost always associated with transcriptional activation (Bannister and Kouzarides, 2011). The presence of acetylated histone residues is maintained by a two-prong system that includes the activation of acetyl transferases (HATs), and, crucially, the inhibition of histone deacetylases (HDACs). Histone acetylation is often seen as the most crucial epigenetic "switch" to allow interconversion of permissive and repressive chromatin states by acetylating and deacetylating key lysine residues (acetylation reading, and writing is summarised in Marmorstein & Zhou, 2014). These proteins, like their methylation counterparts exist in large complexes that are highly conserved (Haberland, Montgomery & Olson, 2011). The changes that acetyl marks make on histones are well explained chemically. In short, these acetyl groups cause the amino acid it is attached to lose a significant portion of its positive charge, weakening the interaction between the histone and the surrounding DNA and thus providing more fluid access to the transcriptional machinery (Hong et al., 1993). A simple model of histone acetylation is provided in Fig. 1.1

**Figure 1.1** Simple model of histone acetylation and deacetylation as an epigenetic switch. This model illustrates shows how acetylation can be used as a switch to compact chromatin and supress gene expression or widen it and make it accessible to the transcriptional machinery. Schematic drawing taken from Verdin & Ott (2015).


Reader
Ac
Ac
Ac
Ac
Acetylated chromatin
Open and transcriptionally active
Gene
On
Writer
Chromatin
remodelling
Eraser
Deacetylated chromatin
Compact and transcriptionally repressed
Gene
Off
Nature Reviews | Molecular Cell Biology

### 1.1.3 Chromatin modifications at regulatory regions

Histone modifications are powerful agents of transcriptional control because they can affect histones in the proximity of not only both gene promoters and coding sequences, but also in proximal or distal enhancer elements. Broad scale meta-studies from the ENCODE project estimate that there are over 1 million enhancers in the human genome (The ENCODE Project Consortium, 2007). This makes them by far the most prominent elements in the DNA sequence outnumbering genes by nearly 50-fold. Enhancer elements contain some distinguishable and some overlapping histone modulations from promoters. H3K4me3 and H3K27Ac are the two most predominating transcriptional activation marks found at promoters of actively transcribed genes (The ENCODE Project Consortium, 2012). However, these marks have also been associated with active enhancers. H3K4me1 is also associated with enhancers; however, these enhancers need not be activated (Zentner & Scacheri, 2012). Indeed H3K4me1 enrichment has be associated with dormant enhancers during development where it has been shown to both precede enhancer activation, as well as persist after enhancer de-activation in some cases (Creyghton et al., 2010; Bonn et al., 2012). Recent work has examined how mono-methylation of H3K4 seems to act as a priming event to ready an enhancer for activation. In this situation enhancers are also marked by H3K27me3 enrichment that can become activated with H3K27Ac deposition (Rada-Iglesias et al., 2010).

The specific tri-versus mono methylation of H3K4 to distinguish enhancers from promoters has also been a target of research. The present molecular explanation for differential levels of methylation at enhancer versus promoters is dependent on the existence of DNA methylation at these sites through the activity of CXXC proteins which read non-

methylated CpG island which can be found at enhancers but are not found at promoters of actively transcribed genes (Sharif-Zarchi et al., 2017). It is thought that methylation at H3K4 sites occurs through a gradient moving from a singular methylated site-through a di methylated site towards the final tri-methylated residue through a series of differing COMPASS complexes (Miller et al., 2001). These sites are also demethylated by unique demethylases specific to a given degree of methylation (Reviewed in Hyun et al., 2017). The degree of methylation at a given region on H3K4 does seem to be dependent on the amount of time these methylation complexes can be bound to the region, and therefore many other molecular mechanisms outside of DNA methylation are probably in play to repel the complexes after mono-methylation at enhancer sites but not at promoter sites.

Broadly speaking, H3K4me1-H3K27Ac abundance relative to H3K4me3-H3K27Ac abundance can be used to estimate if a given sequence belongs to a promoter or an enhancer element with some degree of accuracy (Calo & Wysocka, 2013). H3K4me1 versus H3K4me3 methylation can hence be used to discriminate enhancers from promoters when examining the chromatin profiles of a given sequence independently of any other genomic or epigenetic elements. It is worth noting that many exceptions to the rule exist and several enhancers with abundant H3K4me3 enrichment have been described (Russ et al., 2017).

### 1.1.4 Chromatin dynamics during development of the nervous system

Changes in histone modification during development can occur through several mechanisms. The most obvious is through both global changes in expression of the aforementioned acetyltransferases, deacetylases and methyltransferases. Studies have shown

that, during development of the nervous system, the global levels of H3 acetylation (including principally permissive marks H3K27Ac, H3K9Ac) appear to be stronger in early differentiating neurons than those of more mature migrating neurons (Cho et al., 2011). There are also key tissue, and niche specific events, especially in progenitor and stem cells of the nervous system, that show abundant increases or decreases in histone modifications; both through localised changes in expression of the modifying proteins, and through recruitment-specific events these proteins to the chromatin (reviewed in Lomvardas & Maniatis, 2016).

A great deal of attention has been paid to the changes over time in important individual enhancer and promoter elements of developmental transcription factors. Interestingly, many key genes involved in developmental pathways have also been shown to have coincident enrichment of both permissive (H3K4me3) and repressive (H3K27me3) histone marks. These areas, so called “bivalent” promoters, are emerging as an important developmental regulator to poise the expression of genes kept in repressive states in the absence of differentiation signals, especially in ES cells (Voigt, Tee, & Reinberg, 2013). Some studies have also found bivalent chromatin in neural progenitors at genes that are moderately expressed (Lim et al., 2009)

### 1.2 The nervous system in vertebrates

During development, vertebrates generate a complex network of interconnected cells called neurons. The network of these cells constitutes the whole of the organism’s nervous system and is broadly split into two classes – the central nervous system (CNS), which includes the brain and the spine, and the peripheral nervous system (PNS), which broadly includes all of

the nerves found outside of the brain and spine. The PNS is further subdivided into the somatic nervous system, which comprises both efferent and afferent sensory neurons, and the autonomic nervous system which contains both the sympathetic and parasympathetic systems.

Both CNS and PNS utilise the activity of neurons. There are many ways to categorize these neurons – one of the most common is to label them as sensory neurons, motor neurons or interneurons. Sensory neurons are afferent neurons that carry signal from the PNS to the CNS. Motor neurons are efferent neurons that send a signal from the CNS to the muscle system to initiate a muscle contraction via actin-myosin fibers. Interneurons on the other hand are the most abundant class of neurons in vertebrates and act as mediators between neurons; notably in the CNS. In the CNS, especially in the brain, a signal may go through several mediators before reaching its final destination.

Another useful way of categorising neurons is by organising them by the type of neurotransmitter they release. Neurotransmitters come in several varieties but are broadly thought of as being either excitatory (inducing an excitatory signal in a postsynaptic cell) or inhibitory (block the firing of the postsynaptic cells). The most common neurotransmitter in the vertebrate CNS is glutamate; which is almost always excitatory in nature and is believed to make up approximately 50-80% of the synapses in the brain (Lesley & Rodin, 2011). The next most common is Gama-aminobutyric acid (GABA), which is generally inhibitory, and makes up the majority of the remaining 20% of the synapses in the brain (Lesley & Rodin, 2011). Other transmitters located in the brain include serotonin, noradrenaline, dopamine and acetylcholine which, despite being less prevalent than the abovementioned transmitters in the brain, are vitally important in the proper functioning of the brain and are often linked to the onset of

mental health illnesses. Acetylcholine, which is responsible for inducing muscle contractions, is the most common neurotransmitter in the whole organism (Lesley & Rodin, 2011).

The nervous system also contains another class of cell called glial cells, or simply glia. Glia are believed to be about as abundant as neurons in many organisms, but in many others such as humans, can vastly outnumber them. Glia and neuron numbers are reviewed in von Bartheld, Bahney, & Herculano-Houzel (2016). There has been an increasing interest in the study of glial cells in recent years as the literature has expanded upon their importance during development. The role of glia cells is reviewed by Stevens (2003); I briefly summarise some of these roles here. Not only do glia serve a role in structurally supporting nerves, they also play a significant role in directing cells and programming cellular proliferation and fate commitment by sending out signals to the neurons in their niche. They also provide cues for guiding axon growth to make sure axons can be projected to their proper targets in a complex environment. Finally, glial cells, in the form of oligodendrocytes or Schwann cells, play a significant role in allowing nerve cells to carry signals more efficiently. Schwann cells are specialised cells that coat neuron axons and produce the myelin sheath which is crucial for insulating an AP as it propagates down the axon to ensure the loss of basal signal over time is minimized.

Together, neurons and glia comprise the nervous system, a system that is crucial for vertebrate function. Most importantly, the nervous system is required for communication between organs and tissues in a very precise manner and at great speeds. Over evolution its role has expanded, and it is now responsible for several complex modalities including systems of behaviors and information processing.

## 1.3 Vertebrate CNS development

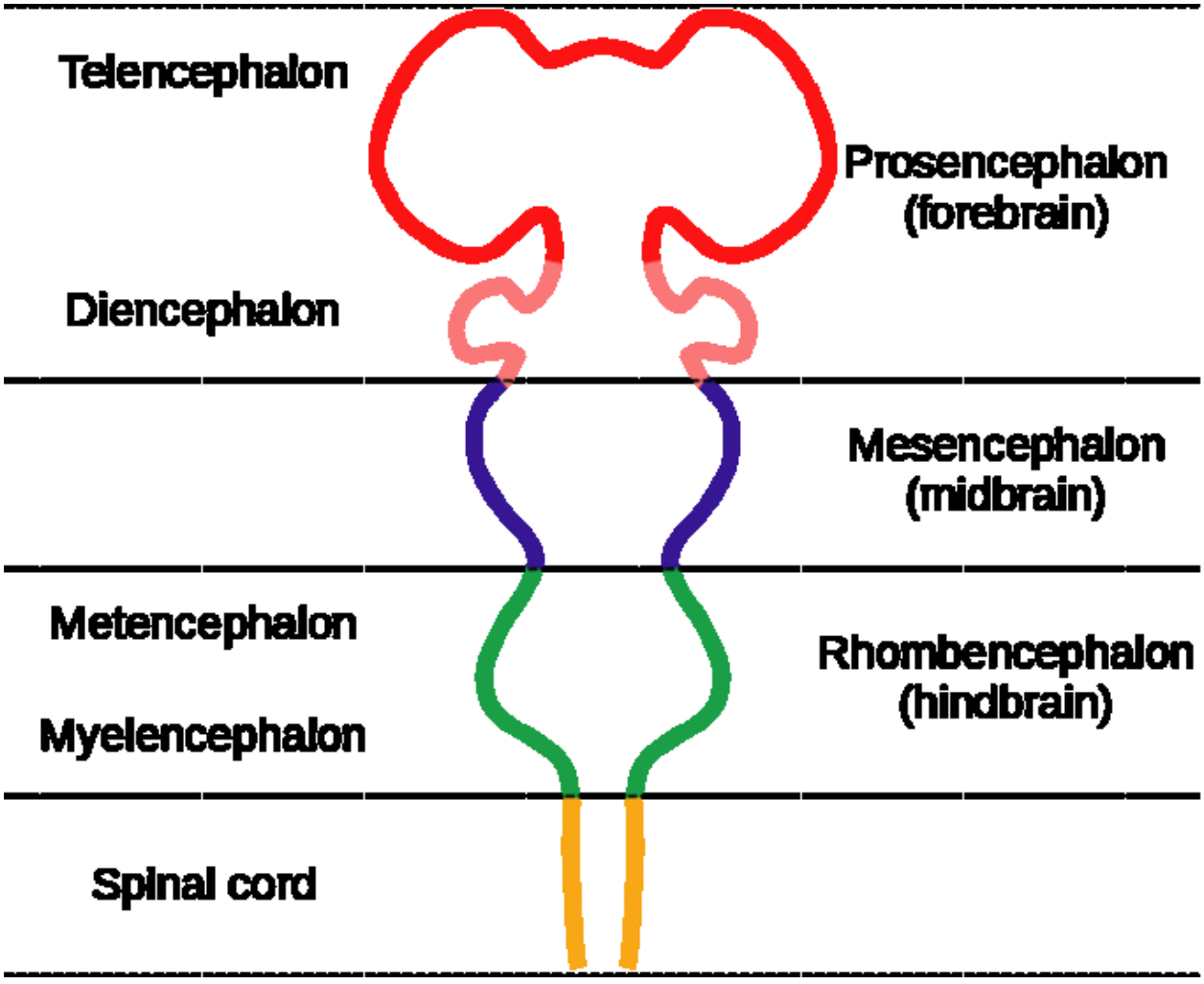
### 1.3.1 Early nervous system development in vertebrates

The development of the nervous system in vertebrates starts with formation of the neural plate. The neural plate is a transitory structure defined by ectodermal cells which align at the centre of the dorsal side of a developing embryo. During embryogenesis, a longitudinal groove of the neural plate deepens and ridges on each side of the plate (neural folds) become elevated before meeting and forming the neural tube. This process of neurulation is extremely complex and involves over 300 genes in mammals (Reviewed in Nikolopoulou et al., 2017). After neurulation, the neural tube pinches off from the surrounding epidermal tissue to form a neural and an epidermal layer. The cells on either side of the roof of the neural tube, called neural crest cells, undergo an epithelial to mesenchymal transition. These cells will migrate away and ultimately become various cell types including those of the PNS.

The neural tube will subsequently differentiate over early embryogenesis into four major areas – the prosencephalon (forebrain), the mesencephalon (midbraind), the rhombencephalon (hindbraind) and the spinal cord. This process is broadly achieved in three ways as summarised by Gilbert (2000). First, anatomical shifts of the neural tube lead to bulges and constrictions that will form the chamber of the proto-brain and spinal cord. Second, cellular populations within the walls of the neural tube rearrange themselves to form a series of progressively restricted and specialised domains. And third, neuroepithelial cells differentiate into the various cell types present in the CNS. These processes are largely driven by gradients of transcription factor expression and morphogens that occur along both an anterior-posterior

and a ventral-dorsal axis. An illustration of the four major zones of the early developing human CNS are shown in Fig. 1.2

**Figure 1.2** Schematic representation of the embryonic vertebrate brain. These regions will further differentiate during embryogenesis into sub structures within each protogroup.


Telencephalon
Diencephalon
Prosencephalon
(forebrain)
Mesencephalon
(midbrain)
Metencephalon
Myelencephalon
Rhombencephalon
(hindbrain)
Spinal cord

### 1.3.2 Vertebrate brain development

The development of the brain is particularly complex due to the many structures that have evolved unique roles over time. After the neural tube gives rise to the forebrain, midbrain, and hindbrain, each of these segments further complexifies during embryogenesis. The forebrain first divides into the telencephalon and the diencephalon; the hindbrain divides into the metencephalon and myelencephalon while the midbrain remains undivided for some period of time during neural development (The molecular pathways governing early brain development regionalisation are reviewed in Kiecker & Lumsden, 2005). The telencephalon and diencephalon further subdivide into much more specialised regions. The Diencephalon notably becomes the thalamus and the hypothalamus.

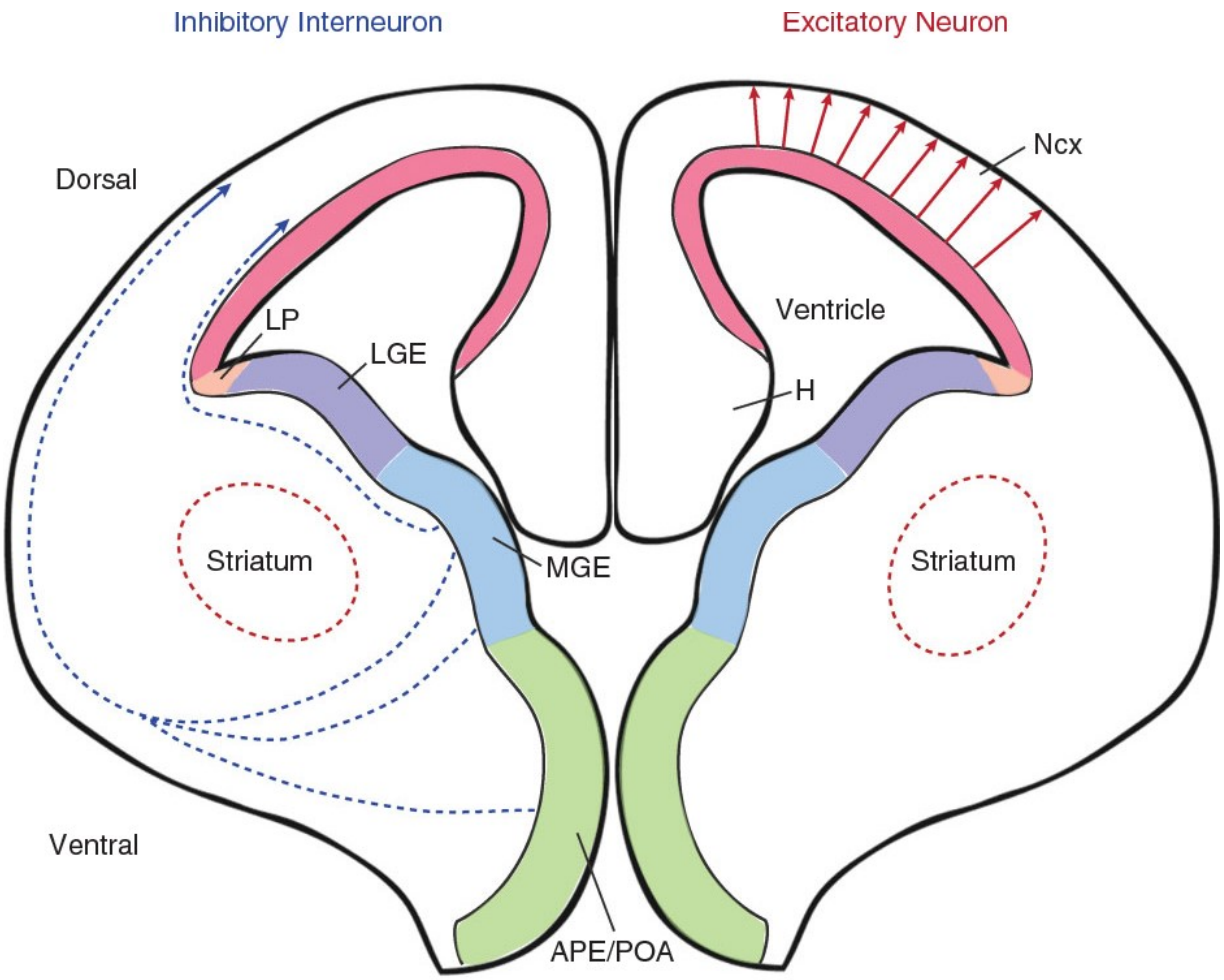
The telencephalon has the most complex embryogenic path because it is responsible for giving rise to several brain regions involved in complex tasks and behaviors with equally as complex neural topologies. The developing telencephalon contains two regions of highly proliferative cells, the dorsal telencephalon and the ventral telencephalon. The dorsal telencephalon expresses transcription factors like *Pax-6* and will give rise to the cerebral cortex and the hippocampus (Caric et al., 1997; Warren et al., 1999). The ventral telencephalon, the region that will be the focus of this dissertation, is marked most notably by expression of Nkx gene family members (Price et al., 1992). Crucial to this study is the expression of *Dlx* transcription factors in the ventral telencephalon during mid gestation (Dollé, Price & Duboule, 1992).

The ventral telencephalon is important for establishing a lot of the interneurons that populate the brain including those found in both the striatum and the cortex. The ventral

telencephalon is known to contain two regions of proliferation. First, is the ventricular zone (VZ) which is only found during embryogenesis. Second is the subventricular zone (SVZ) which persists after embryogenesis and remains a zone of neurogenesis even in the adult brain (neurogenesis in the adult brain SVZ is reviewed in Alvarez-Buylla & García-Verdugo, 2002). In midgestation the ventral telencephalon is anatomically marked by clusters of cells which bulge from the ventricular space called the lateral, medial and caudal ganglionic eminences (LGE, MGE, and CGE respectively) so named based on their location within the SVZ

During telencephalic development neurons become post mitotic within these proliferative zones of the VZ or SVZ or within the dorsal telencephalon. They then migrate towards the mantle zone (MZ) and differentiate. Most of these migrations are radial; as such most neurons within a given region are derived from the most proximal prolferative zone (O'Rourke et al., 1992). However, in some instances, cells migrate tangentially, as is the case of several cortical interneurons. These neurons are born in the VZ as radial glia cells that then migrate to the neocortex during gestation (O'Rourke et al., 1992; O'Rourke et al., 1995). As they migrate, they become more specialised and commit to location dependent fates. These tangentially migrating cells are of special interest in this thesis. A schematic representation of a mouse coronal section indicating the major areas of the telencephalon is presented in Fig. 1.3

**Figure 1.3** Schematic representation of a coronal section of a mouse telencephalon at E13.5. Schematic includes the ventral and dorsal telencephalon regions. The red arrows and dashed arrows indicate the direction of radially migrating neural progenitors, and tangentially migrating neural progenitors respectively. APE= anterior entopeduncular area, H = hypothalamus, LGE = lateral ganglionic eminence, LP = lateral plate, MGE = medial ganglionic eminence, POE = preoptic area.


Inhibitory Interneuron
Excitatory Neuron
Dorsal
Ncx
LP
LGE
Ventricle
H
Striatum
MGE
Striatum
Ventral
APE/POA

### 1.3.3 Tangential migration in the forebrain of mammals

The tangential migration of interneurons is a complex cellular process. Tangential migration is reviewed in Marín and Rubenstein (2001). These neurons move in a path that is in congruent with the brain surface in a matter than defies the conventional boundaries of the developing brain, typically in a MGE-LGE-destination directionality. This migration requires a wide range of guidance cues including SLIT1/2 (Marín et al., 2003). One explanation for how tangential migration occurs involves the activity of another group of cells called “corridor cells” first described by López-Bendito and colleagues (2006). Corridor cells migrate from the LGE to the MGE and embed themselves in the mantle zone. Here they provide a permissive corridor for the cells migrating to the neocortex as well as the growth of thalamocortical axons which extend from the striatum to the neocortex via the release of certain guidance signals, including but perhaps not limited to NRG1.

### 1.3.4 GABAergic neurogenesis in the forebrain of mammals

GABAergic neurons are neurons that release the neurotransmitter GABA and are hence - as previously mentioned - are predominately inhibitory in nature. GABA is a neurotransmitter synthesized by the enzyme glutamic acid decarboxylase (GAD). GABAergic cells directly convert glutamate into GABA using one or two potential GAD isoforms, GAD65 and/or GAD67; often referred to as GAD1 and GAD2 respectively (Kaufman, Houser & Tobin, 1991). Recently a newly discovered GAD3 isoform was discovered in some hominids via sequence synteny (Grone & Maruska., 2016).

The GABAergic interneurons that populate the cortex are among these neurons born in the GEs of the ventral telencephalon. During development these cells migrate and differentiate tangentially towards the neocortex. Indeed, approximately 70% of the total interneurons in the neocortex are derived from the MGE (Miyoshi et al., 2010) Other GABAergic neurons born in the GEs also migrate to other structures including a much shorter tangential migration into the striatum (López-Bendito et al., 2006; Stopczynski, Poloskey & Haber, 2008).

Functionally, the cortical GABAergic interneurons principally play a role of modulating cortical circuitry. Defects in these neurons have been associated with several disorders including autistic spectrum disorders (ASD), schizophrenia and epilepsy among others. These disorders are largely due to electrophysiological impairments in the cortex – in short, a lack of proper signal inhibition. The genetics and function of neocortical GABAergic interneurons is reviewed in Rossignol (2011).

## 1.4 *Dlx* homeobox genes

### 1.4.1 Overview of homeobox genes

Homeobox genes are a group of genes that are responsible for controlling the pattern of body formation during embryogenesis. These genes largely share a DNA binding motif known as a homeodomain, coded by a DNA motif named the homeobox. The homeodomain contains a helix-turn-helix motif that contacts the DNA through the major groove with a 3 base-pair-periodicity (McGinnis et al., 1984b; Fongang et al., 2016). The homeobox motif is approximately 180 bp (around 60aa in the homeodomain) and was first discovered in the Drosophila *HOM-C* genes (McGinnis, et al., 1984a, b; Scott et al., 1984). The DNA target of these proteins largely

contain a tetranucleotide ATTA core-motif downstream of a GG dinucleotide (Kissinger et al., 1990). The homeodomain also contains an N-terminus tail that is believed to drive specificity of the homeodomain to its target DNA sequence (Qian et al., 1994). Homeobox genes are well conserved across the kingdom animalia and are found in species ranging from sponges (Seimiya et al., 1994) to humans (Levine et al., 1984). In general, homeobox genes are responsible for patterning the embryo and specifying the body plan. They play a vital role in development and deletions or mutations in these genes can lead to gross changes in morphology, changes broadly referred to as homeotic transformations (Emerald & Roy, 1997).

Several homeobox genes are grouped into clusters and are often found to have several paralogs with a high degree of sequence identity. The prevailing theory on homeobox genes is that they largely arose from duplication events of a primordial homeobox gene set, explaining the large degree of sequence identity of both coding sequences, and in some cases cis regulatory elements that regulate them. Homeobox gene evolution is reviewed by Duboule (2007). Interestingly, these genes are often organized colinearly, meaning their position on the chromosome 3'-5' correlates with their expression domains along the anterior to posterior (AP) axis (Gaunt 1988). This phenomenon is best observed in the *Hox* genes which contain a very large number of clusters and paralogues and whose expression domains in the anterior to posterior direction are directly proportional to their position within a cluster 3'-5' (Gaunt et al., 1986; Gaunt 1988; Duboule & Dollé, 1989). The evolutionary explanation for the presence and maintenance of collinearity amongst *Hox* genes is not readily transparent but may be a result of widespread sharing of *cis-regulatory* elements as well as chromosomally dependent silencing effects during development occurring in a 3'-5' direction, as proposed by Gaunt (2015).

### 1.4.2 *Dlx* gene overview

The *Dlx* genes of vertebrates are the orthologues of the *Drosophila Distal-less* (*Dll*) genes; genes responsible for specifying distal structures in the developing embryo. In *Drosophila,* the *Dll* gene was so named because loss of its function perturbed proper morphogenesis of the distal units of the organism, including the antenna and the later tarsal segments (Cohen et al., 1989; Dong, Chu & Panganiban 2000). In vertebrates there are 6 *Dlx* genes whose expression is largely restricted to embryogenesis. The profiles of homeobox gene expression in embryonic and adult tissues is examined in Dunwell and Holland (2016). These genes have been implicated in the development of several structures that are broadly distal to the organism including but not limited to: the craniofacial skeleton (Robledo et al., 2002), dentition (Thomas et al., 1997), limbs (Bulfone et al., 1993), and sensory afferents (Kaji & Artinger., 2004). This thesis will focus on the rather abundant expression of these genes in the forebrain (Dollé, Price and Duboule, 1992).

### 1.4.3 *Dlx* function in the telencephalon

In the telencephalon of mammals, 4 of the 6 *Dlx* genes are expressed, specifically - *Dlx1, Dlx2*, *Dlx5* and *Dlx6.* In the GEs, expression of the *Dlx* genes occurs in this order: *Dlx2* → *Dlx1* → *Dlx5* → *Dlx6i* (Eisenstat et al., 1999). *Dlx* expression domains are visualised in Fig. 1.4. Some evidence suggests *Dlx1, Dlx2, Dlx5* and *Dlx6* expression continues into adulthood in the cortex of mammals; however, the role of *Dlx* genes in adulthood has not been as well characterised as their roles in development (Saino-Saito, Berlin & Baker, 2003; Wang et al., 2010).

The majority of what we have learned about *Dlx* genes function in the telencephalon comes from the production of mice which lack one or multiple *Dlx* genes. Mice which lack either *Dlx1 (Dlx1$^{-/-}$)* or *Dlx2* (Dlx2$^{-/-}$) activity have relatively normal forebrains during development (Qiu et al., 1995). This phenomenon has been suggested to be a consequence of functional redundancy of *Dlx1* and *Dlx2* genes which is further given credence by the fact that the two genes are expressed in the same cells during cellular differentiation (Eisenstat et al., 1999). Postnatally, *Dlx1*$^{-/-}$ mice show specific differentiation changes in interneuron populations such as reductions of calretinin$^{+}$ and somatostatin$^{+}$ GABAergic interneurons in the forebrain (Cobos et al., 2005) indicating that *Dlx* genes are important in the development of *GABAergic* interneurons. On the other hand, *Dlx2*$^{-/-}$ mice are more similar to their wildtype (WT) counterparts save for a subset of cranial neural crest cells that are re-specified (Qiu et al., 1995). Double mutant mice lacking both *Dlx1* and *Dlx2* fail to develop GABAergic interneurons in the neocortical, hippocampal and olfactory regions of the brain, but do have neurons migrating to the striatum (Anderson et al., 1997a; Anderson et al., 1997b) These mice also have an accumulation of neurons within the proliferative zones of the ventral telencephalon (Anderson et al., 1997b). Lastly, these mice also display reductions in expression of *Dlx5* suggesting that *Dlx1* and *Dlx2* are potential activators of *Dlx5* (Anderson et al., 1997a).

Removal of *Dlx5* activity is lethal in early postnatal development (Acampora et al., 1999; Depew et al., 1999). A proportion of these mice, about 28%, have exencephaly and as such the brains of these mice are hard to study in late embryogenesis (Depew et al., 1999). Mice without both *Dlx5* and *Dlx6* activity have an even more severe phenotype than mice without *Dlx5* alone (Robledo et al., 2002). These mice have normal telencephalic patterning in early and

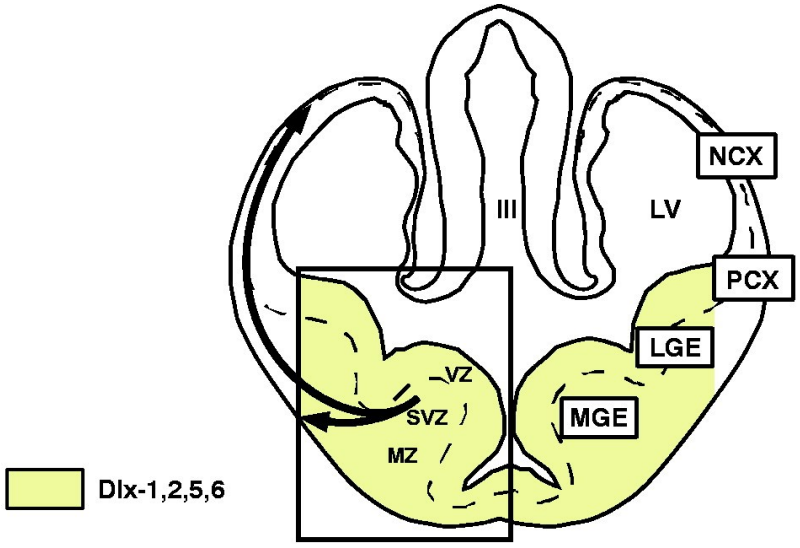
mid gestation, but the migration of some GABAergic interneuron populations is perturbed once again linking *Dlx* activity to GABAergic interneurons. Indeed, these mice display a significant reduction in the number of mature parvalbumin$^{+}$ GABAergic neurons in the neocortex (Wang et al., 2010). These mice ultimately die between E18.5 – P1 due in part to severe exencephaly (Robledo et al., 2002). Moreover, the heterozygous deletion has only a very mild phenotype suggesting dose specific activity of these genes (Robledo et al., 2002).

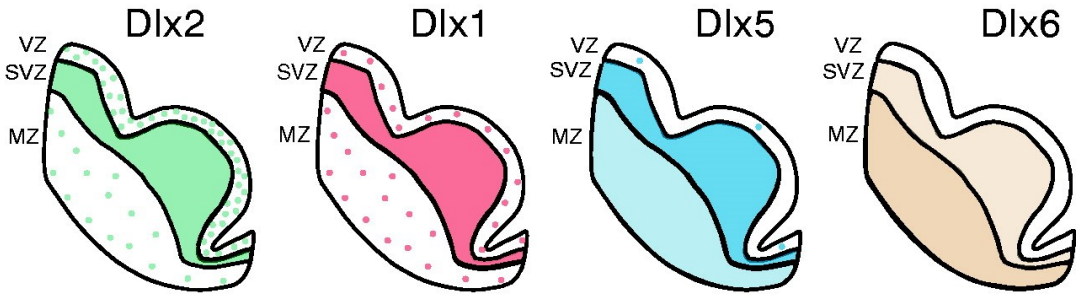
Studies in cultured cells have also been performed to elucidate the role *Dlx5* and *Dlx6* may play in later stage neurogenesis. Cultured neural stem cells (NSC) from newborn mice devoid of *Dlx5* function have been shown to have reduced capacity to generate neurons, but neurons derived from E12.5 embryos do not have any difficulty generating neurons (Perera et al., 2004). Rescue experiments involving forced *Dlx5* overexpression in cultures of newborn NSCs restored neurogenic potential. Together, these results further suggest that *Dlx5* and *Dlx6* play a role in late-gestation and early postnatal neurogenesis from the ventral telencephalon but not during early gestation. Earlier neurogenesis from the ventral telencephalon may be more dependent on *Dlx1* and *Dlx2* activity, a model that is sensible given the earlier activation of *Dlx1 and Dlx2.*

In addition to the information gleaned from removing the activity of *Dlx* genes, as well as *Dlx* expression in NSCs, several pieces of evidence in the literature suggest that *Dlx* genes are vital for the establishment of the GABAergic identity of a large number of interneurons. Firstly, the *Dlx* expression domains overlap with those of *Gad* and *Mash1* (*AscI1)* linking their expression spatially with the first GABAergic progenitors in the GEs (Porteus et al., 1994). It has been shown that *Dlx5* and *Dlx2* are able to induce ectopic GABA expression in coronal brain

slices (Stühmer et al., 2002a). More interestingly it has also been demonstrated that *Dlx* genes, specifically *Dlx1* and *Dlx2*, directly activate the transcription of both *Gad* isoforms in GABAergic progenitors (Le et al., 2017).

**Figure 1.4** *Dlx* gene expression in the E12.5 mouse telencephalon (Panganiban & Rubenstein, 2002).




Model:

**Dlx2 → Dlx1 → Dlx5 → Dlx6**

### 1.4.4 *Dlx* chromosomal organization

The *Dlx* gene structure is well reviewed by both Panganiban & Rubenstein (2002). Some of the key points are highlighted below. *Dlx* genes in vertebrates are chromosomally organized in bigene clusters. In mice and humans, there are 3 clusters – *Dlx-1/2, Dlx-5/6* and *Dlx-3/4* (previously *Dlx7/8*). Each cluster is on a discrete chromosome linked to a corresponding *Hox* gene cluster; for instance, the *Dlx1/2* cluster located on chromosome 2 in mice and humans is syntenic to the *Hoxd* cluster (McGuinness et al., 1996). Each *Dlx* cluster is composed of two convergently transcribed genes, meaning the 5' to 3' transcription of the two genes is convergent. The *Dlx* genes of mice and humans and contains a common structure – 3 exons and 2 introns, with the homeodomain being split between the 2$^{nd}$ and 3$^{rd}$ exon (Price et al., 1991; McGuinness et al., 1996; Ellies et al., 1997; Liu et al., 1997). Each cluster contains a central intergenic region of approximately 8-10Kb in length. These intergenic regions are known to contain some important cis regulatory elements, or CREs (Zerucha et al., 2000; Ghanem et al., 2003). Each intergenic element is approximately 250-450bp in length.

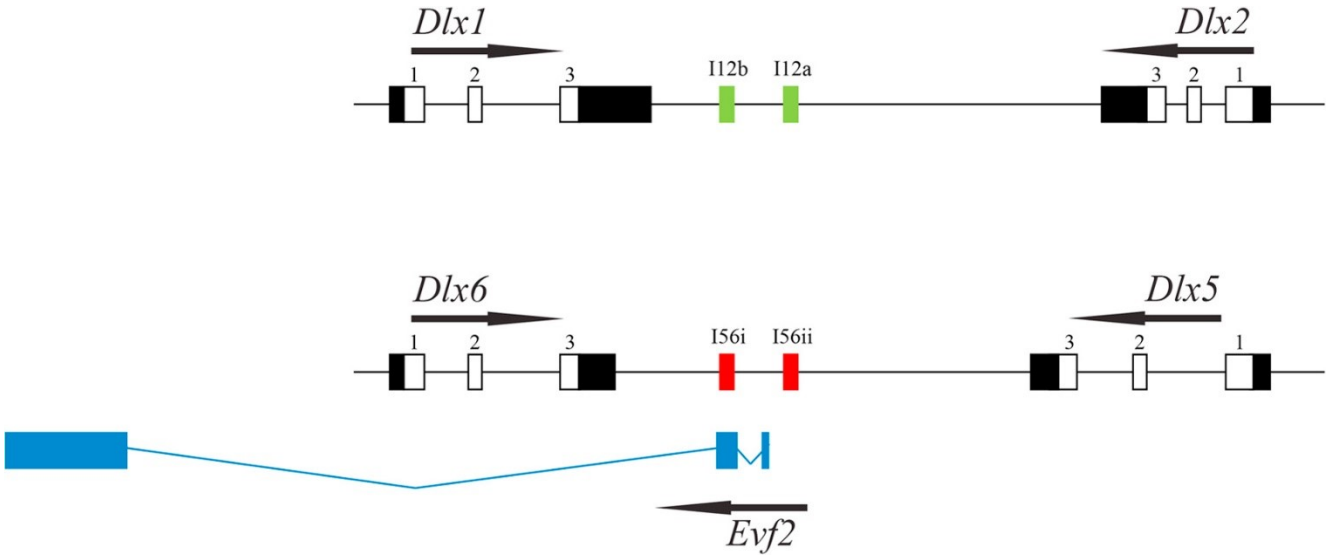
The *Dlx* genes contain a high degree of sequence conservation between species, especially in the homeodomain. More interestingly, there is also a high degree of conservation between the intergenic enhancers of various species (Zerucha et al., 2000) There is however, very little similarity between enhancers within the same species). Much like general *Hox* gene evolution, the current model of explaining *Dlx* gene evolution is based upon a tandem duplication to produce a cluster followed by duplication events tied to the generation of the four *Hox* clusters, preceded by steady divergent evolution (Stock et al., 1996). This evolutionary

model helps explain why the bigene clusters are similar between organisms but not between each other.

Since this thesis focuses on *Dlx* genes expressed in the forebrain, the organization of the *Dlx5/6* and *Dlx1/2* bigene clusters is of special interest here. The *Dlx1/2* bigene cluster is marked by the presence of two intergenic enhancers I12a, and I12b; the latter of which has been implicated in *Dlx* expression in the forebrain along with another regulatory element upstream of *Dlx2* called URE2 (Ghanem et al., 2003). The *Dlx5/6* bigene cluster also contains two intergenic enhancers named I56i, and I56ii which both have forebrain activity.

There is also the presence of the *Evf2* gene, coding for a lncRNA, and whose promoter co-localizes with the I56ii enhancer. Interestingly *Evf2* lncRNA has been shown to directly interact with *Dlx2* and, in doing so, increases transcriptional activity of the *Dlx5/6* enhancers (Feng et al., 2006). *Evf2* also been shown to have chromatin remodelling inhibition activity in complexes with Dlx1 and Brg1, as well as acetylation activity to increase acetylation levels at H3K9, H3K18, and H4K5 at the I56i and I56ii enhancers (Cajigas et al., 2015). *Evf2* also affects enhancer methylation activity through Mecp2 recruitment to methylate the I56i region (Berghoff et al., 2013). A schematic of the *Dlx1/2* and *Dlx5/6* loci is presented in Fig. 1.5.

**Figure 1.5** Schematic representation of the *Dlx5/6* and *Dlx1/2* bigene clusters in the mouse. Note the presence of the *Dlx* intron exon-structure and the location drawn to scale of the intergenic enhancers. lncRNA *Evf2* whose promoter overlaps with the I56ii enhancer, and whose second exon is coincident with the I56i enhancer is also drawn. (Taken from Darbandi et al., 2016)

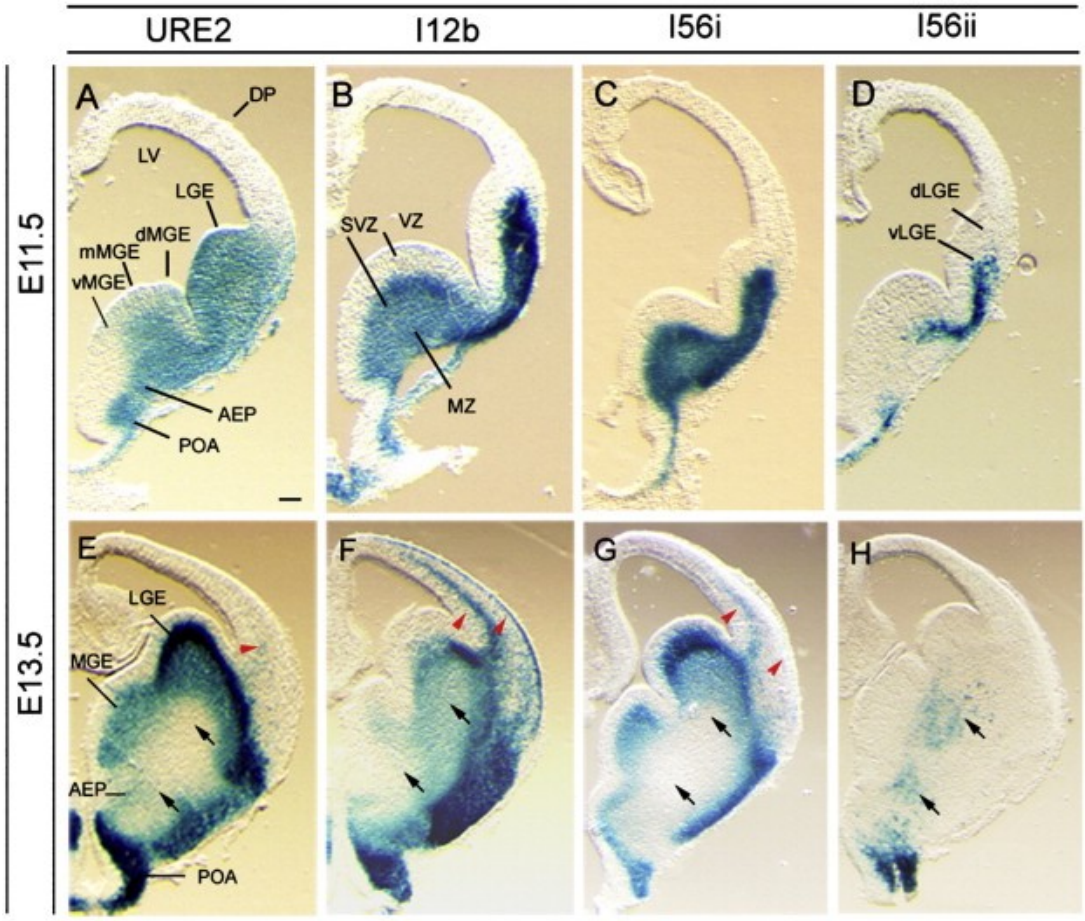

Dlx1
Dlx2
1
2
3
I12b
I12a
3
2
1
Dlx6
Dlx5
1
2
3
I56i
I56ii
3
2
1
Evf2

### 1.4.5 *Dlx* regulation through forebrain enhancers

The *Dlx* genes are regulated in a complex fashion by various factors that can interact with *Dlx* promoters and/or their CREs. As previously mentioned, the picture is made more complicated by the Dlx proteins ability to activate other *Dlx* genes in a time dependent fashion (Zerucha et al., 1997). The intergenic *Dlx* enhancers have unique forebrain activity wherein they drive *Dlx* expression in distinct but sometimes overlapping domains of the developing forebrain during mid-gestation (Ghanem et al., 2003; Ghanem et al., 2007; Zerucha et al, 2000). It is possible that these cis regulatory elements are capable of driving expression of both flanking genes, and this may help explain some degree of *Dlx* functional redundancy although this has never been conclusively demonstrated.

In the *Dlx1/2* bigene cluster, I12b is active in the forebrain, specifically in the mantle zone (MZ) and the sub-ventricular zone (SVZ) of the MGE and LGE as demonstrated by reporter lines wherein mice carry a transgene driven by the I12b element fused to a beta-globin minimal promoter (Ghanem et al., 2003). URE2 also has activity in the forebrain within the VZ, the SVZ and the MZ of both LGE and MGEs. I56i has activity in a similar fashion as I12b but presumably drives expression of a different set of *Dlx* genes. I56ii activity on the other hand is restricted mostly to the deep MZ and the striatum, especially later in gestation, indicating that I56ii corresponds to *Dlx* expression in the striatum and not in the neocortex (Ghanem et al., 2003; Ghanem et al., 2008). A summary of these findings is provided in Fig. 1.5. Our lab further demonstrated that these enhancers ultimately label discrete populations of cells in the adult cortex (Ghanem et al., 2007).

**Figure 1.6** Activity of *Dlx* forebrain enhancers illustrated by *LacZ* reporter constructs. Staining is performed on coronal slices of E11.5 and E13.5 transgenic mice. (Adapted from Ghanem, 2007).


URE2
I12b
I56i
I56ii
E11.5
E13.5
A
B
C
D
E
F
G
H
DP
LV
LGE
dMGE
mMGE
vMGE
AEP
POA
SVZ
VZ
MZ
dLGE
vLGE
MGE

Our laboratory has gained a lot of insight into *Dlx* regulation via the deletion of enhancers I56i, I56ii, and I12b. Deletion of each one of these enhancers results in mild phenotypes, and in all cases homozygous mice are fertile, viable, and display no gross changes in anatomy or morphology. I56i null mice displayed reduced expression of *Dlx5* and *Dlx6*, slightly increased expression of *Dlx1* and normal levels of *Dlx2* compared to their WT littermates (Esau, 2013). These mice also show significant reductions in *Gad* expression, but we were unable to detect any significant changes in GABAergic subtype populations at P35 (Esau, 2013; Darbandi 2014; Zhou, 2017). I56ii$^{-/-}$ mice show a similar trend in expression patterns and exhibit impaired expression of striatal markers like *Islet1* further solidifying I56ii's role in directing *Dlx* expression to the striatum (Fazel Darbandi et al., 2016). I12b$^{-/-}$ null mice, on the other hand, have moderately reduced expression of *Dlx1* and *Dlx2* but retain relatively normal levels of *Dlx5, Dlx6*, and G*ad* suggesting that I12b may not contribute to driving expression of these genes in the ventral telencephalon as significantly as another regulatory region, like URE2 (Yu, 2011).

Our lab has also had an interest in understanding if the *Dlx* CREs and promoters were able to crosstalk with each other or interact with CREs or promoters on another *Dlx* cluster in trans. We have postulated that these interactions may go beyond the sharing of RNA Pol II occupancy between an enhancer and the promoter it is proximal to, and this may further explain some degree of how *Dlx1/2* expression can affect *Dlx5/6* expression. To examine this we preformed some chromosome confirmation capture (3C) experiments that are able to detect if there are 3D interactions occurring between regulatory elements. While the data is limited, we do find some evidence of interactions between the *Dlx1* promoter and CRE I56i in

E13.5 ventral telencephalon suggesting that the transcription of *Dlx1* could be linked to I56i activity despite its traditional association with *Dlx6*

### 1.4.6 *Dlx* regulation through transcriptional activators

Our group and others have determined various activators of *Dlx* genes. One such gene is *Mash1* (*Ascl1*) which has been demonstrated to activate *Dlx1/2* expression via interactions with the I12b enhancer (Poitras et al., 2007). The important developmental signaling molecule Sonic hedgehog (Shh) has also been shown to be involved in the activation of *Dlx2*, *Dlx5* and *Evf2* expression in the forebrain (Feng et al., 2006). We have also shown that the general transcription factor 2i (Gtf2i) binds to the I56i enhancer in an ultraconserved region (Poitras et al., 2010). SNPs in this region dramatically reduce Gtf2i binding. Interestingly GTF2I is hemi-deleted in human patients who suffer from a rare neurodegenerative condition called Williams Syndrome (WS). We have shown that mice with I56i deletions, and more strikingly with I56i/I12b double deletions show several behavioral phenotypes that correspond to WS syndrome including increased fear and anxiety and deficits in learning/memory linking the importance of *Dlx* genes to neurodevelopmental disorders (Zhao, 2017).

## 2 Statement of inquiry

The regulation of *Dlx* genes has been studied extensively at the level of the genome; however, very little work has been done to understand how *Dlx* genes may be regulated at the level of the chromatin. We postulate that *Dlx* CREs likely play a multifaceted role in regulating *Dlx* genes by conferring specific chromatin profiles to the genes that flank them. These effects may be especially significant because the CREs are so close to the genes and thus the enhancer and promoter profiles of the clusters may influence each other.

Since *Dlx* genes exhibit complex regulatory networks, we expect a relatively dynamic chromatin landscape with permissive chromatin marks in tissues that express *Dlx*. We believe that the establishment of an epigenetic signature that pervades the bigene clusters may help explain how *Dlx* expression is fine-tuned in the ventral telencephalon.

To address this question, I used both *in silico* (ENCODE) and *in vivo* (ChIP-qPCR) methods to understand the chromatin landscapes of the *Dlx1/2* and *Dlx5/6* bigene clusters at various points in development. I studied the chromatin landscapes of the *Dlx* clusters during midgestation in ganglionic eminences isolates when the *Dlx* genes are robustly expressed, and a high degree of tangential migration is occurring. To understand what effects the enhancers may play on conferring epigenetic regulation, I studied wildtype mice and compared them to mice which have targeted deletions of the I56i, I56ii and I12b enhancers. I examined some of the most abundantly enriched histone modifiers corresponding to enhancer and promoter elements – namely; H3K4me1, H3K4me3, H3K27Ac and H3K27me3 to observe what histone residues were modified, and if these CREs were important in maintaining these profiles. Finally,

I used qRT-PCR to tease out genotype and development specific changes in some key histone modifying enzymes to understand if changes in *Dlx* expression levels correlate with changes in expression of these histone modifier genes, or if changes in enrichment levels of modified histones are a consequence of localization or recruitment effects.

# 3 Materials and methods

## 3.1 Mouse husbandry

### 3.1.1 Housing and breeding

Adult mice of various genotypes were housed in ventilated cages with siblings of the same sex not to exceed 4 mice per cage. Mice were fed a standard diet of basic pellets and were given water bottles that were replaced once a week. The mouse facility lighting is on an automatic schedule of a 12 hour light/dark schedule and is maintained at 25°C. For breeding; mice selected were between 9 weeks – 26 weeks of age and did not share any parents. All protocols used in this study were approved by the University of Ottawa's Animal Care Ethics Community. The mice are further looked after by members of the University of Ottawa's ACVS to check for general health and mortalities. Fighting or problematic mice are separated and placed in individual cages when the need arises.

### 3.1.2 Weaning and ear-tagging

Mouse pups were weaned at approximately 21-24 days postnatal. Individuals of the same sex with the same parentage were housed in the same cage up to a maximum of 4 mice per cage. To mark the mice, each weaned mouse had its ear-tagged and was given a unique number so its genotype could be tracked over-time. To genotype juvenile mice, a small piece (about 3-5mm) of the distal tip of the tail was removed and placed in a 1.5mL Eppendorf tube 1 week after weaning. The tail piece was stored at -20°C until genotyping was completed (see 3.4).

### 3.2 Generation of mutant mice

Mice with targeted deletions of I56i, I56ii and I12b were generated at the Transgenic Mouse Core Facility at McGill University using homologous recombination in Embryonic Stem (ES) cells. The production of recombined ES cells was done in the Ekker laboratory by former post-doctoral fellow Luc Poitras. To generate these cells, he used homologous recombination (HR) in a Bacterial Artificial Chromosome (BAC) that contained the mouse sequence of the *Dlx* cluster and its respective intergenic enhancers. HR was used to replace the intergenic enhancers with a floxed PGK-neomycin resistant cassette. Successful recombination was screened using a neomycin plate and confirmed via sequencing. Positively recombinant BACs were transformed into mouse ES cells and screened. Positive cells were cultured and injected into a blastocyst which was subsequently implanted into a pseudo-pregnant C57BL/6 mouse. Some of the pups of this litter were chimeric. Founder mice were identified using PCR and mated with Sox-CRE mice to remove the neomycin cassette. These mice were outbred with C57BL/6 WT mice to isogenize the background. Heterozygous mutant mice were mated together to produce homozygous mice.

### 3.3 Dissection of ganglionic eminences and tails

GEs from E11.5, E13.5, and E14.5 mouse embryos were dissected in a similar fashion. First a mating pair of a female and male mouse was generated in the early afternoon. Typically, mice crossed were heterozygous to produce both WT and homozygous offspring. After establishing the cross, female mice were checked each subsequent morning for a vaginal plug.

The appearance of a vaginal plug was taken as E0.5. Mice were dissected on the appropriate embryonic day at approximately the same time of day regardless of age. Female mice that were pregnant were euthanized with $CO_2$ followed by dislocation of the cervical vertebrae. To remove the embryos, the female was cut in the abdomen along the midline to expose the uterus. The uterus was then removed using dissecting scissors and the embryos were placed in a 15mL tube containing ice cold PBS.

The embryos were subsequently removed from the uterus using scissors. At this point the amniotic sac is torn and the embryos are separated from the placenta using fine tweezers. Each embryo was then placed in an individual labelled well of a 6 well plate containing ice cold PBS. A small piece of the tail is taken for genotyping (See 3.4). At this point, the embryos are decapitated and each head is placed in a small Petri dish for finer dissection.

To remove the GEs, the brain was carefully removed from the brain case using tweezers under a dissecting microscope (Nikon; SMZ1000). The isolated brain was then cut with scissors to isolate the ventral telencephalon. MGE, and LGEs were isolated together using the tweezers to peel off the neocortical and striatal tissues surrounding them. Where applicable, as at later stages of development (E14.5), the CGEs which is caudal to the fusion point of the LGEs and MGEs was also extracted. Isolated GEs were put into 1.5mL tubes containing PBS and centrifugated for 5 minutes at 4°C and 1500 rpm in a Sorvall™ Legend™ Micro 21R Microcentrifuge (Thermo Scientifc™). The supernatant was discarded, and the tissue was flash frozen in liquid nitrogen. Tissue was stored at -80°C where it could be used for either ChIP or RNA analysis.

### 3.4 Genotyping

Tail tissue from embryonic mice, or from adult mice was genotyped using the Phire Animal Tissue Direct PCR Kit® (Thermo Scientific™) according to the manufacturer's information. PCR conditions were as follows: 98°C for 5m; followed by 40 cycles of 98°C for 5s, 58-63°C for 5s and 72°C for 25s; followed by a final extension step of 72°C for 1minute. PCR reactions were run on a standard 1% agarose gel against GeneRuler DNA Ladder Mix (Thermo Scientific™). The gel was stained with RedSafe (FroggaBio) and visualized under UV light using AlphaImager HP system (ProteinSimple). The following primers were used to genotype:

DI56i F - CAGTTCTAAGCAGAGTTCTAG, DI56i R – CTCAGTCAGTCTTCAGAATGG,

DI56ii F - ACGGAAGCAAGACAGGCAAG, DI56ii R - GAGGTGGCTTTGGTGGAGAG,

DI12b F – GGCAAAATGCAATTTTGGGA, DI12b R - GTCTGAAGAACCATCTAACCTG

### 3.5 RNA analysis

#### 3.5.1 RNA extraction

RNA was extracted from dissected GE tissue using the RNEasy Mini Kit® (Qiagen) with some modifications. Embryonic tissue was homogenised by vigorous pipetting in 800µL of RLT Plus Buffer substituted with 1/100 volume β-mercaptoethanol. It was then passed through the columns according to the manufacturer's instructions. RNA was eluted using 20µL RNAse free water pre-warmed to 37°C and quantified using a NanoDrop™ 2000c (Thermo Scientific™). RNA integrity was analysed by running 300ng of RNA on a standard 1% agarose gel against

GeneRuler DNA Ladder Mix (Thermo Scientific™). The Gel was stained with RedSafe (FroggaBio) and visualised under UV light AlphaImager HP system (ProteinSimple). Only RNA samples where the 28S band was as intense or greater then the 18S band were used for cDNA synthesis

### 3.5.2 cDNA synthesis and qRT-PCR

cDNA was generated using 500ng RNA using the iScript™ cDNA Synthesis Kit (Bio-rad) according to manufacturer's information. Oligonucleotide primers were validated to meet MIQE guidelines. For a list of primers see Table 3.1. qRT-PCR was run using 5µL 2X PowerUp™ SYBR® Green Master Mix (Applies Biosystems) 0.5µL F/R primer (10mM), 0.5µL $H_2O$ and 4µL cDNA. The reaction was run on an Eco™ Real-Time (Illumina Inc) machine using the following conditions: 50°C for 2m, 95°C for 2m; followed by 40 cycles of 95°C 15s, 60°C for 1m. qRT-PCR Reactions were run in experimental duplicates.

### 3.5.3 Statistical analysis and data presentation of qRT-PCR data

To generate fold expression values, data was transformed using the Pfaffl method (Pfaffl, 2001). For normalization, target genes were compared to the geometric mean of two housekeeping genes – *β-Actin* and *GapdH* which were determined to be experimentally stable genes between genotypes. Data is presented as mean with error bars corresponding to SD of the $\Delta\Delta C_q$ values. Statistical analysis was preformed on $\Delta C_q$ values using a two-tailed t-test and an n of 3. Traditional p values demarcations were used wherein: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Multiple comparisons were corrected for using the Holm-Sidak method. Calculations were preformed on GraphPad - Prism8 software.

**Table 3.1 –** Complete list of primers used for qRT-PCR experiments. Primers were sourced from the literature and were experimentally determined to meet MIQE guidelines.

| Primer Name | Sequence (5' to 3') |
|---|---|
| Ezh1 F | AAAGTCAACACTTCCCGCTG |
| Ezh1 R | CATACAGAGCCTTTGCTCCCA |
| Ezh2 F | ACTGCTTCCTACATCCCTTCC |
| Ezh2 R | GTGCTGGGTCTGCTACTGTT |
| Ash1l F | GGGCACGAGAAGGATGATGA |
| Ash1l R | TCATGGGTACCTCCCTGTCC |
| Kmt2c F | TGTGAACAAGGGTTCCCGAG |
| Kmt2c R | GGGAGACAGTGCACATCCAA |
| Setd5 F | GCGTATCACCACTGACCCAA |
| Setd5 R | ATCCAGCGCTTCTTACAGGA |
| Chd8 F | GACCGAGGAGCAGGTTCAAA |
| Chd8 R | TACGGGTGATTGCAGCACTT |
| Hdac1 F | TTACTACTACGACGGGGATG |
| Hdac1 R | ACCATAGTTGAGCAGCAAAT |
| Hdac2 F | ATGGCGTACAGTCAAGGAG |
| Hdac2 R | CCATTTTTCGGTATAAACCAT |
| Hdac3 F | ATCATGCCAAGAAATTTGAG |
| Hdac3 R | AGGTAGAAGGCTTCCTGAAC |
| Kat2a F | GGCTTCTCCGCGAATGACAA |
| Kat2a R | GTTTGGACGCAGCATCTGGA |
| Kat2b F | GACACCAAACAAGTCTATTTCTACCTC |
| Kat2b R | ACCACTGGCCGTGTCATCT |
| Gapdh F | TGCCCCCATGTTTGTGATG |
| GapdH R | TGTGGTCATGAGCCCTTCC |
| B-Act F | CTGTCCCTGTATGCCTCTG |
| B-Act R | ATGTCACGCACGATTTCC |

### 3.6 *In silico* ChIP-Seq analysis

Data from Roadmap Epigenomics Project and ENCODE was mined through the ENCODE portal database and visualized through the *UCSC* genome browser using the mouse mm10 (GRCm38) build. Images were generated as PDF files through the browser and edited appropriately on Adobe Illustrator as advised in the portal. To normalize the data sets, maximum enrichment levels were set to be the same between all developmental time points of a given enriched histone. Data were presented as mean data with no error (whisker) bars for clarity. Data are from the most stable release from the Ren Lab that was published in 2016 and represents relatively limited (2X) coverage.

### 3.7 ChIP-qPCR analysis

#### 3.7.1 Isolation of nuclei and histone immunoprecipitation

Native ChIP was carried out similarly to like Magklara et al. (2011) with several modifications. Tissue was homogenised by vigorous pipetting in 200µL Buffer I (0.3M Sucrose, 60mM KCl, 15mM NaCl, 5mM $MgCl_2$, 0.1mM EGTA, 15mM Tris-HCl pH 7.5) substituted with 0.5mM DTT, 0.1mM PMSF, 1 X Halt™ Protease Inhibitor Cocktail (Thermo Scientific™), and additionally 50mM $Na(C_3H_7COO)$ when using α-acetyl antibodies. The homogenate was centrifugated at 1500xg for 5mins at 4°C. The pellet was resuspended gently in 100µL Buffer I, then mixed with 100µL of Buffer II (0.3M Sucrose, 60mM KCl, 15mM NaCl, 5mM $MgCl_2$, 0.1mM EGTA, 15mM Tris-HCl pH 7.5, 0.4% NP-40) substituted identically as Buffer I. Lysis is carried out on ice for 7 mins. The lysed cells are then pelleted at 8500rpm for 7mins at 4°C. Nuclei are gently resuspended in 200µL Micrococcal nuclease (MNase) buffer (0.32M Sucrose, 50mM Tris-

HCl pH .5, 4mM $MgCl_2$, 1mM $CaCl_2$) substituted with 0.1mM PMSF and 1 X Protease Inhibitor Cocktail, and 50mM Na($C_3H_7COO$) where necessary . Nuclei are allowed to sit on ice for 30 minutes – 2 hours to facilitate resuspension. Nuclei were counted using a hemocytometer at a 1:25 dilution in a 0.2% Trypan Blue. Micrococcal nuclease digestion was completed using 1µL of undiluted enzyme MNase (NEB) at 37°C for 3-5mins to produce a population of mono-tetra nucleosomes (S1 fraction). At E13.5, and E14.5 500,000 nuclei/IP was digested, at E11.5 a whole pair of GEs (250,000-400,000 nuclei) was digested. Digested chromatin is mixed with 1/20$^{th}$ volume 0.5M EDTA and centrifuged for 10mins at 10,000rpm and 4°C. The supernatant is decanted and placed in a clean 1.5mL tube at 4°C overnight (S1 fraction) The pelleted material, which is heterochromatin is resuspended in 200µL dialysis buffer (1mM Tris-HCl pH 7.5, 0.2mM EDTA) substituted identically to MNase buffer and further digested overnight at 4°C on a slow rocker.

The next day the heterochromatin sample is centrifuged for 10mins at 10,000rpm and 4°C. The supernatant is drawn off (S2 fraction) and combined with the S1 fraction. The volume in the tube is subsequently adjusted to 700µL with Wash Buffer I (50mM Tris-HCl pH 7.5, 10mM EDTA. 125mM NaCl, 0.2% Tween, 5mM Na($C_3H_7COO$)) substituted with 1X Protease Inhibitors and 20mM Na($C_3H_7COO$) when necessary. Chromatin fractions were next incubated with 10µL Protein A and 10µL Protein G Dynabeads (Invitrogen Life Sciences) that have been resuspended in 100µL Wash Buffer I/IP at 4°C for one hour to pre-clear the chromatin. After, the chromatin is removed from the beads and topped up to 1mL with Wash Buffer 1 substituted as above. At this point, 100µL of the chromatin is put aside as the input fraction (10%). The rest of the chromatin was incubated overnight at 4°C with one of: 1.5µL of anti-H3K4me3 (07-473;

Millipore), 2.5µL of anti-H3K27Ac (07-360; Millipore), 2.5µL anti-H3K27me3 (07-449; Millipore), 2.5µL anti-H3K9Ac (ab10810; Abcam), 5µL anti-H3K4me1 (710795; Invitrogen Life Sciences), or 0.75-2.5µL anti-green fluorescent protein (A-11122; Invitrogen Life Technologies). During this time 10µL Protein A and 10µL Protein G Dynabeads/IP are resuspended in 1mL of Blocking Buffer (Wash Buffer I substituted with 0.3mg/mL yeast t-RNA, 0.2% BSA) and blocked overnight at 4°C with gentle rocking.

The following day, the blocked beads are resuspended in Wash Buffer I (100µL/IP) substituted as above. 100µL of the blocked beads in Wash Buffer I is added to each IP and incubated at 4°C for 3-4 hours with gentle rocking. The samples are then washed for 5 mins each in the following buffers: 4 X Wash Buffer I, 3 X Wash Buffer II (50mM Tris-HCl pH 7.5, 10mM EDTA, 175mM NaCl, 0.1% NP-40, 5mM $Na(C_3H_7COO)$), and 1 X TE pH 8.0). DNA is stripped from the beads by incubating in 100µL Elution Buffer (10mM Tris-HCl pH 8.0, 1mM EDTA, 100mM $NaHCO_3$, 1% SDS) for 15 mins then removing the solution from the beads. This is done 2 times for a total volume of 200µL. The solution is then incubated at 65°C for 30 mins to destroy any remaining proteins. Concurrently, the 10% input fraction is thawed and adjusted to 200µL in Elution Buffer. It is also digested at 65°C for 30 mins. ChIP DNA is purified using the ChIP DNA Clean and Concentrator Kit (Zymo Research) according to manufacturers information and eluted with 25µL (IP sample) or 250µL (Input fraction). ChIP DNA is diluted 4-fold in $ddH_2O$ before proceeding to ChIP-qPCR

### 3.7.2 qPCR on ChIP DNA

ChIP DNA is analysed using the Eco Real-Time instrument. Primers were validated to meet MIQE guidelines and checked for specificity via running the products on a 2% agarose gel. Primers used for the various loci are listed in Table 3.2**.** qPCR was run with 5µL 2X PowerUp™ SYBR® Green Master Mix, 0.5µL F/R Primer (10mM), 0.5µL $H_2O$ and 4µL ChIP DNA using the following conditions: 50°C for 2 minutes, 95°C for 2 minutes; followed by 40 cycles of 95°C for 15 seconds, and 60°C for 1 minute. qRT-PCR Reactions were run in experimental duplicates.

### 3.7.3 Statistics and data presentation on ChIP-qPCR data

Percentage input was calculated in the following way:

$$100 * (X^{(Cq\ Input - (Log_X 100)) - Cq\ Sample})$$

where X equals the efficiency adjusted primer constant. % Input data were analyzed using a two-tailed t-test comparing mutant to WT data for each line individually. Multiple comparisons were corrected for using the Holm-Sidak method. Data presented is mean values, and error bars displayed on graphs are standard error for each data set. N ≥ 2 for all data sets. Precise n values are listed in each corresponding figure caption. Calculations were preformed on GraphPad - Prism8 software. Traditional p values demarcations were used wherein: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

**Table 3.2 –** Complete list of primers used in ChIP-qPCR experiments. B-Actin and Hoxc10I2 were used as positive and negative control primer sets, changing depending on the precise ChIP experiment preformed. Primers were validated to ensure they met MIQE guidelines.

| Primer Name | Sequence (5' to 3') |
|---|---|
| Dlx6 P F | TAGGGCCACTGCTTCTGAAC |
| Dlx6 P R | TGGAGACTCGGAATAGCCTG |
| Dlx5 P F | GAGAGCTGAGCAGCCTTTTG |
| Dlx5 P R | TTCACGTTTGCAGGAATCTG |
| I56i F | ATTTTTCATGTAGCCCGCTG |
| I56i R | TTGTTTCGCCTTTCCTGTTC |
| I56ii F | ATTGTTTGCACACCCCAGCA |
| I56ii R | CAAAGATGCGCCAGGTCTTGA |
| Dlx 1 P F | TGTTTTGGGGTGTGGAGTGCC |
| Dlx1 P R | ATCACTTACATGGGCAAGGCC |
| I12b F | CGGCCATTGACAAGCTTGCT |
| I12b R | CGCACTGAAAGCCAAATTTGCT |
| Dlx2 P F | CCTTTGGGCTTATAGAGCTCTTCT |
| Dlx2 P R | GCTTCCCAAACTACCCACAGAGAT |
| B-Act F | GAATGTGGCTGCAAAGAGTCTAC |
| B-Act R | CTTCGCTCTCTCGTGGCTAGTA |
| Hoxc10I2 F | GGAGTGAGGCATTGGTTGGA |
| Hoxc10I2 R | CCCAGTACAGCAAAGGGTCC |

## 4. Results

### 4.1 *Dlx*1/2 and *Dlx5/6* clusters have relatively stable epigenetic landscapes in the mouse forebrain during development.

To understand how the *Dlx1/2* and *Dlx5/6* chromatin landscapes were shaped in the developing mouse forebrain, we accessed ENCODE Data from the Ren Lab (2016) and visualised them on the UCSC genome browser. We choose to analyse an approximately 30Kb region that includes the *Dlx* genes, the intergenic regions and a stretch of about 10Kb upstream of each gene to gain a broader sense of how the epigenetic profiles of these loci were situated. We examined forebrain from WT mice at E11.5, E13.5 and E14.5 These are key developmental stages during which the ventral telencephalon is producing new progenitor cells that will make their way towards the neocortex, olfactory bulb, thalamus or striatum. We also examined P0 mice to see if the developmental landscapes of the brain during embryogenesis were maintained in the postnatal forebrain. We examined the histone marks associated with active chromatin promoters, H3K27Ac, H3K4me3, as well as active/putative enhancers via H3K4me1. We also examined H3K27me3 to see if Polycomb repression is involved in silencing *Dlx* expression in the forebrain during development. We chose to focus our attention to the I12a, I56i and I56ii enhancers because we have better characterised these enhancers in the context of forebrain *Dlx* expression.

Firstly, we examined H3K27Ac enrichment which is generally associated with active chromatin throughout the bigene clusters (Fig. 4.1A, B)**.** During development, in the whole mouse forebrain, H3K27Ac is abundant with high peaks associated with the gene exons and

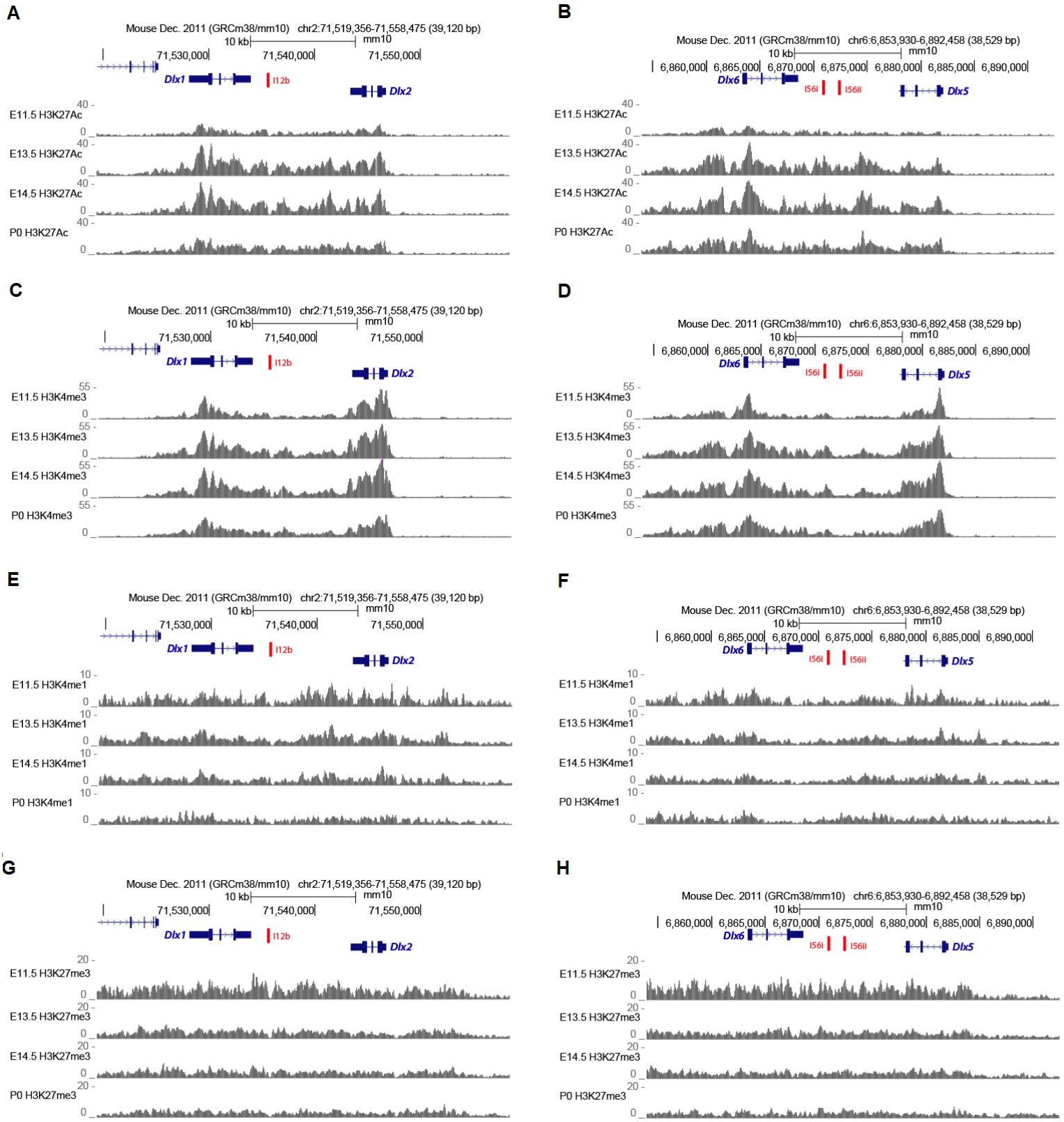
slightly smaller peaks at the enhancers. When comparing developmental time points, we find that the signal is greatest at E14.5, and its lowest at E11.5 indicating expression of the *Dlx* genes may be either more highly expressed at this time point, or that the expression domains make up a larger section of the forebrain. We also note how the enrichment continues upstream of the *Dlx6* promoters but not the *Dlx1/2* and *Dlx5* promoters. This is likely a consequence of the lncRNA *Evf-2* that is expressed in the forebrain and that has exons upstream of the *Dlx6* promoter (Fig 1.5). H3K27Ac is found in the P0 forebrain as well in a similar distribution pattern as during midgestation indicating the persistent expression of all the *Dlx* genes.

The permissive H3K4me3 mark (Fig. 4.1C, D) displays a similar profile to that of H3K27Ac. This mark reaches a slightly higher maximum intensity than the H3K27Ac mark. H3K4me3 is more restricted to the gene exons and the intergenic enhancers than the rest of the intergenic regions. Areas in the intergenic region not corresponding to CREs therefore show lower basal signal than H3K27Ac, whose peaks are a little more broadly distributed across the whole cluster. H3K4me3 changes little during development but has its maximum peaks at E11.5 and bottoms out at P0 in this data set.

H3K4me1 (Fig. 4.1E, F) - a mark associated with enhancer elements, is not highly enriched throughout either bigene cluster. A low level of enrichment (about 4-5X lower than H3K27Ac and H3K4me3 peaks) pervades the entire locus but it does not show obvious peaks that correspond to the CREs or to *Dlx* promoters. In this sense H3K4me1 does not demarcate enhancers in the forebrain. Its enrichment changes very little during development and is found at a low level both during midgestation and in the postnatal forebrain.

Finally, H3K27me3 (Fig 4.1G, H), which corresponds to transcriptional silencing, is lowly enriched throughout the bigene clusters. Much like H3K4me1, its enrichment in the whole mouse forebrain is persistent throughout the locus displaying no discrimination between gene, enhancer, and intergenic elements at any studied time point. It has a maximum signal at E11.5 and decreases during mid-gestation towards birth. The enrichment is about 2X than that of H3K4me1 and about 2X lower than H3K4me3. Interestingly, at E11.5 H3K27me3 enrichment is approximately as abundant as for H3K27Ac, suggesting the existence of a heterogenous cell population where *Dlx* expression is expressed in one tissue and repressed in another. These results also indicate the possibility of a bivalent region of chromatin where H3K27me3 and H3K4me3 are both enriched.

**Figure 4.1** *In silico* ChIP-Seq of the whole mouse forebrain at E11.5, E13.5, E14.5 and P0. H3K27Ac enrichment is displayed for the *Dlx1/2* (A) and *Dlx5/6* (B) clusters. H3K4me3 enrichment is displayed for the *Dlx1/2* (C) and *Dlx5/6 (D)* clusters. H3K4me1 enrichment is displayed for *Dlx1/2* (E) and *Dlx5/6* (F) clusters. H3K27me3 enrichment is displayed for *Dlx1/2* (G) and *Dlx5/6* (H) clusters. Data was acquired from the ENCODE database and visualised on the UCSC genome browser. Data represent published results from the Ren Lab (2016) mapped to the most recent mouse genome release (mm10). The data represent at least 2X coverage and were constructed from libraries containing generally good complexity and little bottlenecking. Enrichment level scales on the Y axis of each plot. Data presented are average enrichment. We omitted the whiskers (SD) for clarity. The location of the relevant *Dlx* forebrain enhancers was annotated manually.


A
Mouse Dec. 2011 (GRCm38/mm10) chr2:71,519,356-71,558,475 (39,120 bp)
10 kb
mm10
71,530,000
71,540,000
71,550,000
Dlx1
I12b
Dlx2
E11.5 H3K27Ac
E13.5 H3K27Ac
E14.5 H3K27Ac
P0 H3K27Ac
B
Mouse Dec. 2011 (GRCm38/mm10) chr6:6,853,930-6,892,458 (38,529 bp)
10 kb
mm10
6,860,000
6,865,000
6,870,000
6,875,000
6,880,000
6,885,000
6,890,000
Dlx6
I56I
I56II
Dlx5
E11.5 H3K27Ac
E13.5 H3K27Ac
E14.5 H3K27Ac
P0 H3K27Ac
C
Mouse Dec. 2011 (GRCm38/mm10) chr2:71,519,356-71,558,475 (39,120 bp)
10 kb
mm10
71,530,000
71,540,000
71,550,000
Dlx1
I12b
Dlx2
E11.5 H3K4me3
E13.5 H3K4me3
E14.5 H3K4me3
P0 H3K4me3
D
Mouse Dec. 2011 (GRCm38/mm10) chr6:6,853,930-6,892,458 (38,529 bp)
10 kb
mm10
6,860,000
6,865,000
6,870,000
6,875,000
6,880,000
6,885,000
6,890,000
Dlx6
I56I
I56II
Dlx5
E11.5 H3K4me3
E13.5 H3K4me3
E14.5 H3K4me3
P0 H3K4me3
E
Mouse Dec. 2011 (GRCm38/mm10) chr2:71,519,356-71,558,475 (39,120 bp)
10 kb
mm10
71,530,000
71,540,000
71,550,000
Dlx1
I12b
Dlx2
E11.5 H3K4me1
E13.5 H3K4me1
E14.5 H3K4me1
P0 H3K4me1
F
Mouse Dec. 2011 (GRCm38/mm10) chr6:6,853,930-6,892,458 (38,529 bp)
10 kb
mm10
6,860,000
6,865,000
6,870,000
6,875,000
6,880,000
6,885,000
6,890,000
Dlx6
I56I
I56II
Dlx5
E11.5 H3K4me1
E13.5 H3K4me1
E14.5 H3K4me1
P0 H3K4me1
G
Mouse Dec. 2011 (GRCm38/mm10) chr2:71,519,356-71,558,475 (39,120 bp)
10 kb
mm10
71,530,000
71,540,000
71,550,000
Dlx1
I12b
Dlx2
E11.5 H3K27me3
E13.5 H3K27me3
E14.5 H3K27me3
P0 H3K27me3
H
Mouse Dec. 2011 (GRCm38/mm10) chr6:6,853,930-6,892,458 (38,529 bp)
10 kb
mm10
6,860,000
6,865,000
6,870,000
6,875,000
6,880,000
6,885,000
6,890,000
Dlx6
I56I
I56II
Dlx5
E11.5 H3K27me3
E13.5 H3K27me3
E14.5 H3K27me3
P0 H3K27me3

**4.2 *Dlx* epigenetic landscapes in the ganglionic eminences are dynamic during development and are characterised by the existence of bivalent chromatin**

In order to get a better sense of how *Dlx* histone modifications change during midgestation, we decided to perform our own ChIP-qPCR experiments focusing on the ventral telencephalon instead of the whole mouse forebrain. In this way, we hope to eliminate some masking effects produced by studying the whole forebrain which contains many non *Dlx* expressing tissues. We once again focused on the widely utilized modifications H3K27Ac, H3K4me3, H3K4me1, and H3K27me3. We focused on areas corresponding to the centre of the CREs as well as areas located approximately 300-800bp upstream of the first exon for each *Dlx* gene, that is, in the approximate regions of the *Dlx* promoters. The location of each targeted area is schematised in Fig 4.2.

We focused our studies on mid-gestation between time points E11.5-E14.5 when the *Dlx* genes are highly expressed in the GEs and when tangential migration is occurring to direct progenitor neurons to either the cortex or the striatum. Examining H3K27Ac enrichment (Fig. 4.3A), we found that a low level of enrichment was relatively consistent between E11.5-E13.5 at all the assayed areas but increased at the *Dlx5/6* cluster significantly at enhancer I56ii and at the *Dlx6* promoter, and not significantly at I56i. All of these regions do correspond with the location of *Evf2* which may explain why *Dlx5* promoter is unaffected. There is also a non-significant increase in signal at the *Dlx1* promoter at E14.5 relative to E13.5. but no changes to *Dlx2* and the I12b enhancer. Changes in signal at I56ii and Dlx6 P approach 300% which seem like significant gains. Despite increases over development the basal signal of H3K27Ac never approach's the same degree of abundance as H3K4me3 or H3K27me3 at each location at E14.5.

H3K4me3 enrichment (Fig. 4.3B) shows this same trend in increases to enrichment at E14.5 compared to E11.5 and E13.5. While these increases fail to reach significance, they still correspond to 2-3X increases across the bigene clusters and represent a much larger raw value. Of the two active chromatin marks, H3K4me3 signal is much greater than H3K27Ac signal at all locations in the clusters and at all points during development.

In line with the *in-silico* findings, we also find that H3K4me1 enrichment is of very low abundance compared to H3K4me3 throughout both bigene clusters (Fig. 4.3C). Indeed, at most times H3K4me1 levels are only slightly higher than the negative control region corresponding to an intron of the *Hoxc10* gene. Moreover, for the most apart, save for one exception, the H3K4me1 mark fails to meaningfully distinguish enhancers from promoters. The one exception is a curious but non-significant increase at the I56ii region at E13.5 of about 300% which subsequently returns to approximately E11.5 levels just one day later. The local spike of H3K4me1 at this time could be a consequence of high I56i activity between E11.5-E13.5 in the GEs.

Curiously, H3K27me3 enrichment increases significantly at all points across the loci between E13.5-E14.5 but not between E11.5-E13.5. This H3K27me3 trend then corresponds to the same trend observed in H3K4me3, that is H3K4me3 also increases at E14. This trend is known as bivalent chromatin. In fact, if one examines more closely the ratio of K4/K27 modified chromatin we find that the increases at E14.5 slightly favour the H3K27me3. Despite H3K4me3 predominating at E14.5 its relative abundance compared to K27 is marginally reduced. It is not readily transparent if changes to the ratio are immediately significant or if this trend corresponds to any changes in expression that are occurring at E14.5 versus E13.5. When

comparing the *in-silico* data set to this one there is a discrepancy in the developmental trend. The whole forebrain data indicates a general decrease in H3K27me3 but increase in H3K4me3 indicating a larger K4/K27 ratio. Again, we cannot preclude the possibility of masking effects due to heterogenous tissue sources making the direct comparison of these two data sets limited.

**Figure 4.2.** Graphical representation of the targeted regions of the *Dlx1/2* and *Dlx5/6* bigene clusters**.** Each arrow point to the centre of a region of approximately 80-140bp region that was amplified via qPCR. We chose regions in the direct centre of the CRE or a segment 300-800bp upstream of the first exon of the closest gene to represent the gene's promoter.

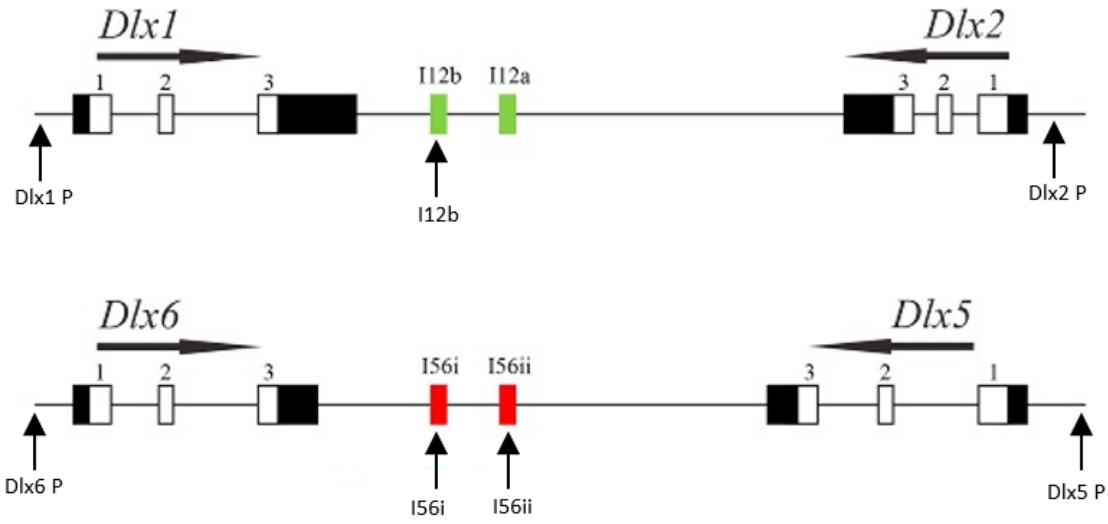
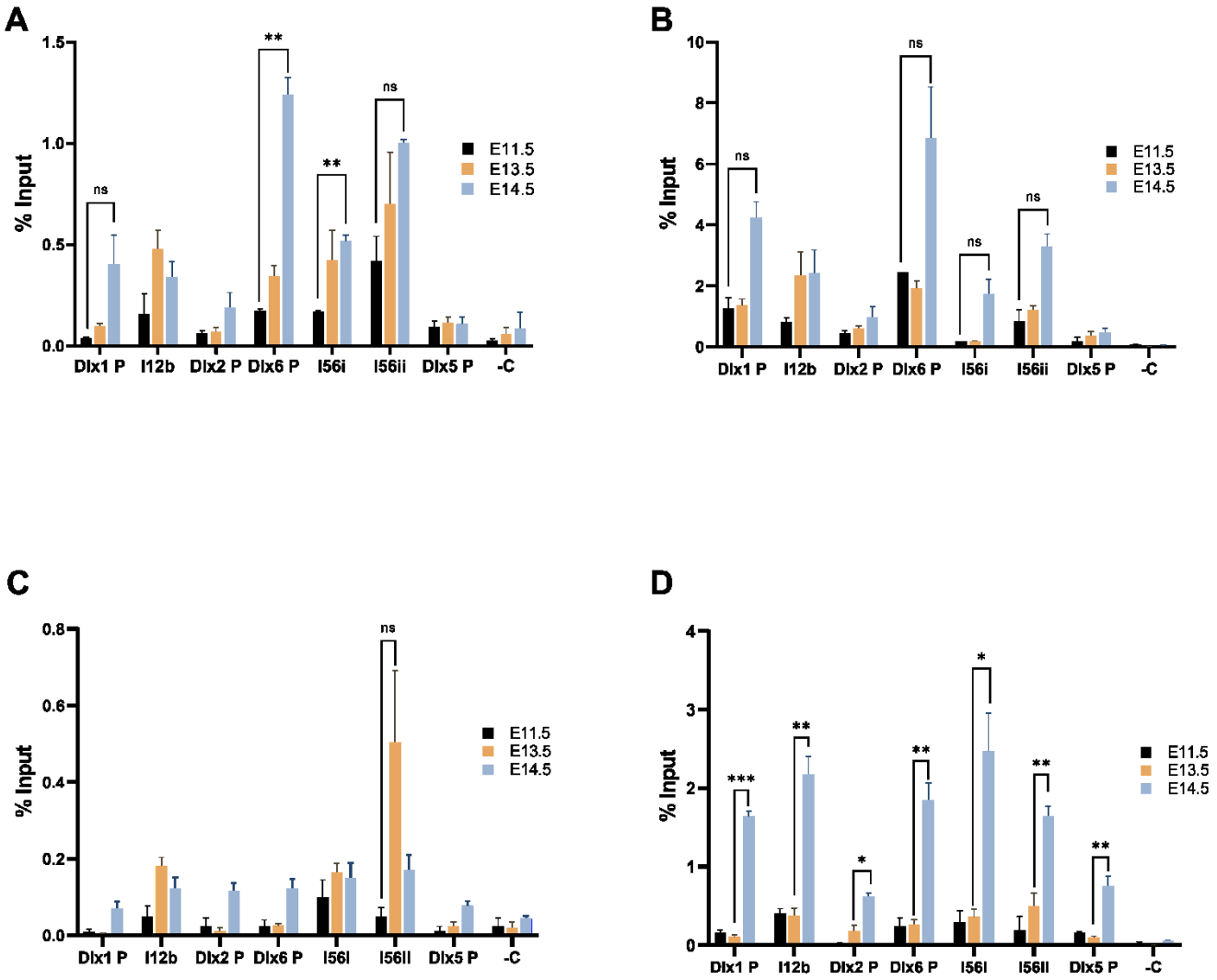

Dlx1
Dlx2
1
2
3
I12b
I12a
3
2
1
Dlx1 P
I12b
Dlx2 P
Dlx6
Dlx5
1
2
3
I56i
I56ii
3
2
1
Dlx6 P
I56i
I56ii
Dlx5 P

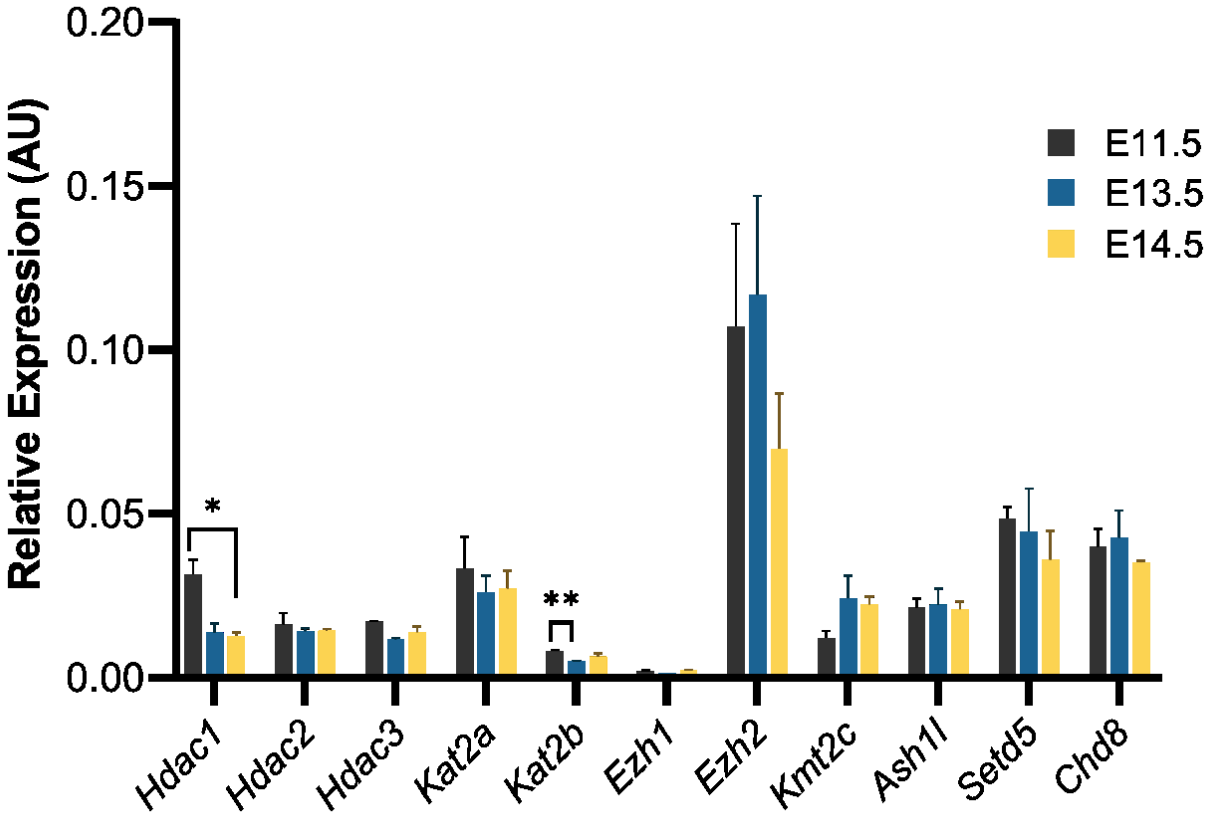
**Figure 4.3.** Histone enrichment at the *Dlx5/6* and *Dlx1/2* bigene clusters in the mouse GE during midgestation**.** ChIP-qPCR Data corresponding to H3K27Ac A), H3K4me3 B), H3K4me1 C) and H3K27me3 D) at E11.5 (black), E13.5 (brown), and E14.5 (blue). Data are presented as % Input using a 1% input control (mean ± SE). Negative control is an intron of *Hoxc10* for all histones except H3K27me3 which used the *β-actin* promoter as a negative control. N ≥ 3 for A, C, D and = 2 for B. Statistical results were calculated using a two tailed t-test on % input values. ns = non significant, * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$


A
% Input
E11.5
E13.5
E14.5
ns
**
**
ns
Dlx1 P
I12b
Dlx2 P
Dlx6 P
I56i
I56ii
Dlx5 P
-C
B
% Input
E11.5
E13.5
E14.5
ns
ns
ns
ns
Dlx1 P
I12b
Dlx2 P
Dlx6 P
I56i
I56ii
Dlx5 P
-C
C
% Input
E11.5
E13.5
E14.5
ns
Dlx1 P
I12b
Dlx2 P
Dlx6 P
I56I
I56II
Dlx5 P
-C
D
% Input
E11.5
E13.5
E14.5
***
**
*
**
*
**
**
Dlx1 P
I12b
Dlx2 P
Dlx6 P
I56I
I56II
Dlx5 P
-C

**4.3 During midgestation the levels of most histone modifying enzymes and chromatin re-modellers are static in the GEs.**

We next sought to understand if any of the changes in histone enrichment observed during midgestation were a consequence of recruitment-specific effects or of changes in tissue-specific levels of histone modifying enzymes over time. To do this we performed qRT-PCR on RNA isolated from the ganglionic eminences at E11.5, E13.5 and E14.5 in WT mice (Fig. 4.4). We found that transcript levels for one histone deacetylase, *Hdac1*, decrease by approximately 50% ($p = 0.31$) between E11.5 and E13.5 and remain at this level at E14.5. Acetyltransferase *Kat2b* also experiences a decrease in expression by nearly 40% specifically between E11.5 and E13.5 ($p = 0.005$), but expression levels go back up at E14.5 to approximately the same levels observed at E11.5. Decreases in expression of both a histone deacetylase and a histone acetyltransferase would be broadly antagonistic; and subsequently it seems likely that changes in modified histone enrichments during development are not a consequence of tissue level changes in expression of the modifying enzymes, but rather of changes in the recruitment of the modifiers to the chromatin. Expression of various histone modifiers including acetyltransferases *Kat2a*, Polycomb group proteins *Ezh1* and *Ezh2*, methyltransferases *Kmt2c*, *Ash1l*, *Setd5*, and the histone remodelling gene *Chd8* remain roughly equivalent in the GEs during midgestaiton. We note that there are many more potential modifiers that we did not target including tissue specific modifiers, and thus we cannot exclude the possibility that an untargeted methyltransferase, for example, is being upregulated between E11.5-E14.5 in the GEs.

**Figure 4.4** Changes in histone modifying enzyme expression during midgestaiton in the GEs**.** qRT-PCR was preformed on isolated GEs from E11.5 (black), E13.5 (blue) and E14.5 (yellow) WT mice. Expression values are given as relative expression to a geometric mean of *GapDH* and *β-Actin* and presented as mean ± SD. N = 3. Statistical analysis was done via a two tailed t-test on $\Delta C_q$ values using Holm-Sidak multiple comparisons method. * = $p \leq 0.05$, ** = $p \leq 0.01$


Relative Expression (AU)
0.20
0.15
0.10
0.05
0.00
E11.5
E13.5
E14.5
*
**
Hdac1
Hdac2
Hdac3
Kat2a
Kat2b
Ezh1
Ezh2
Kmt2c
Ash1l
Setd5
Chd8

### 4.4 CREs confer developmentally specific histone profiles to the *Dlx* bigene cluster in the GEs

In order to understand how the *Dlx* CREs may shape the chromatin profiles of the *Dlx* clusters in the forebrain we preformed ChIP-qPCR experiments during midgestation on mice that had a targeted deletion of on of their Intergenic CREs and compared them to their WT littermates.

We first examined H3K27Ac (Fig. 4.5A), and H3K4me3 (Fig. 4.5B) activating marks in E11.5 embryos. At this time the *Dlx* genes are broadly expressed in WT mice, but significantly altered in the CRE mutant mice our lab has characterised (Fazal Darbandi, 2016; Esau, 2013). Some of the changes we have observed in mice are genotype specific, but broadly these mice experience changes in their *Dlx* expression, as well as in the expression of some downstream targets. Accordingly, we examined the GEs of mice that were missing I56i (ΔI56i), I56ii (ΔI56ii) and I12b (ΔI12b) and compared each of these mutants to the WT and in doing so provide a genotype specific relationship between each CRE and the native chromatin state.

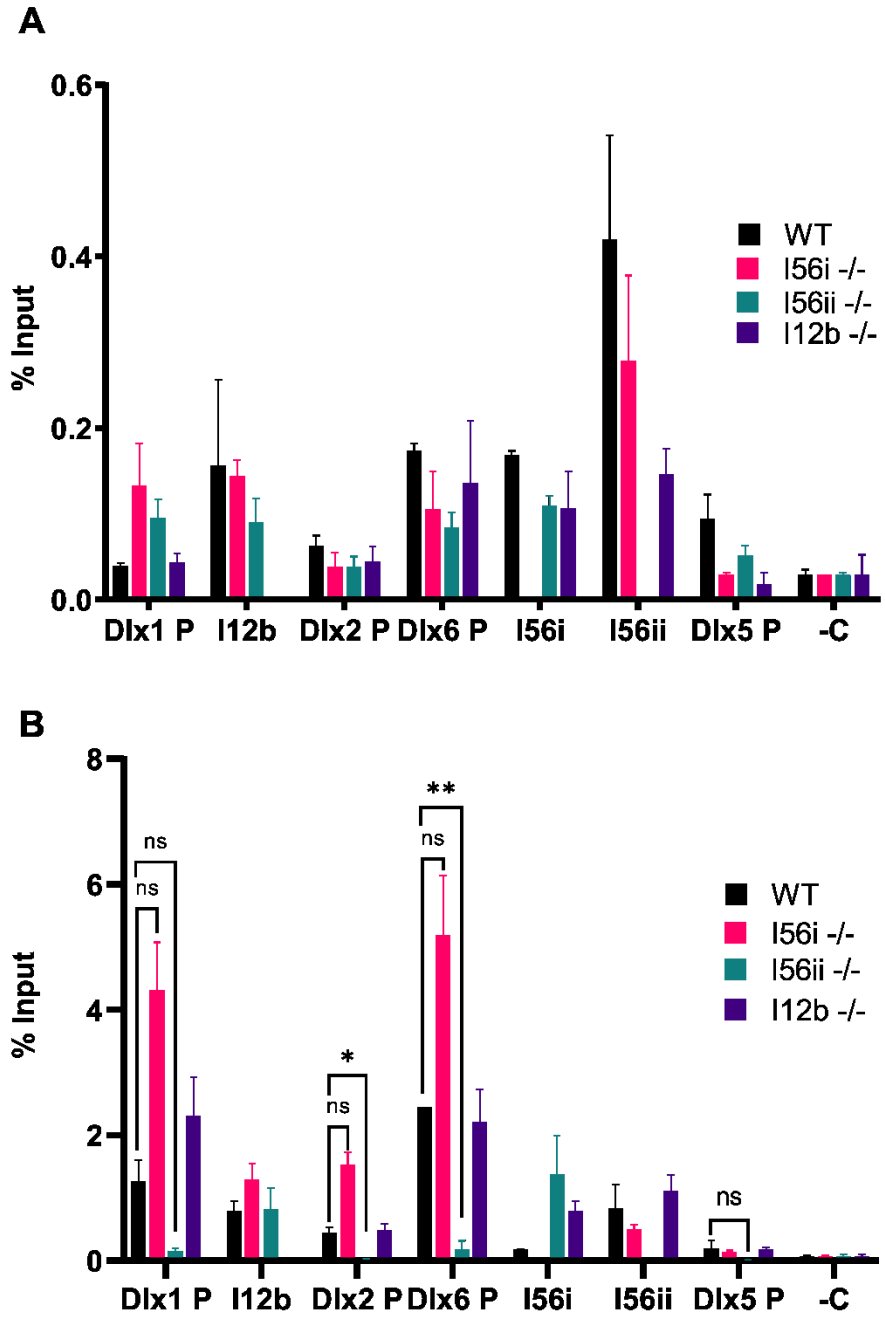
We found that, at E11.5, deletion of the intergenic CREs did not significantly alter enrichment of H3K27Ac at any point in the locus. Levels of H3K27Ac mark remained extremely low relative to H3K4me3 indicating that at E11.5 disruption to the CREs does not affect *Dlx* H3K27 acetylation. (Fig. 4.5A).

On the other hand, some changes in H3K4me3 are observed across the different genotypes. Deletion of enhancer I56i leads to mild increases in H3K4me3 enrichment at the *Dlx1*, *Dlx2, and Dlx6* promoters by about 200%. However, these increases fail to reach significance. Meanwhile, deletion of I56ii leads to an opposite trend with decreases in H3K4me3

at the promoters of *Dlx1* (ns), *Dlx2* (p = 0.042)*, Dlx6* (p = 0.006) and *Dlx5* (ns). Interestingly, ΔI12b H3K4me3 levels are unchanged from the WT levels. Decreases in H3K4me3 correlate with the decreases in expression of *Dlx5* and *Dlx6* in ΔI56ii mice but not with the increase in *Dlx1* or stasis of *Dlx2* expression in these mice (Fazal Darbandi et al., 2016). Likewise, the mild increases in H3K4me3 in ΔI56i mice correlate with increases in *Dlx1* but not with the decreases observed in *Dlx5* and *Dlx6* expression (Esau, 2013).

Taken together, the changes to chromatin observed at E11.5 that differ between all three genotypes suggest that the precise levels of H3K4me3 do not directly correlate with expression levels of these genes. It is possible other activating chromatin marks, like HK36me3 which is deposited in the wake of RNA Pol II binding may change their profiles in a manner that does reflect the transcriptomic changes, but we have not teased this out. Instead it seems as the CREs are important in establishing a stable epigenetic signature and the consequence of removing enhancer activity is a broader change to chromatin signature than a simple reflection of expression dynamics.

**Figure 4.5** Histone enrichment at E11.5 in WT versus *Dlx* CRE mutant mice**.** ChIP-qPCR of E11.5 ventral telencephalon for H3K27Ac A) and H3K4me3 B) on WT (black), ΔI56i (pink) ΔI56ii (green) and ΔI12b (purple) mice. Data are presented as %Input and normalised to the negative control region (Hoxc10I2) Data are presented as mean ± SE, n ≥ 2. Statistical analysis was done via a two tailed t-test on normalized %Input values using Holm-Sidak multiple comparisons method. * = $p \leq 0.05$, ** = $p \leq 0.01$


A
0.6
0.4
0.2
0.0
% Input
WT
I56i -/-
I56ii -/-
I12b -/-
Dlx1 P
I12b
Dlx2 P
Dlx6 P
I56i
I56ii
Dlx5 P
-C
B
8
6
4
2
0
% Input
ns
ns
*
ns
**
ns
ns
WT
I56i -/-
I56ii -/-
I12b -/-
Dlx1 P
I12b
Dlx2 P
Dlx6 P
I56i
I56ii
Dlx5 P
-C

Next, we once again looked at histone modifications H3K27Ac (Fig. 4.6A), H3K4me3 (Fig. 4.6B) in E13.5 mice, that is, slightly later into gestation, to see how alterations in CRE mutant mice have changed at this stage of development when neurogenesis in these regions is occurring. We also looked at H3K4me1 (Fig. 4.6C) enrichment due to increases in H3K4me1 enrichment at I56i and I12b in E13.5 embryos versus E11.5 WT embryos.

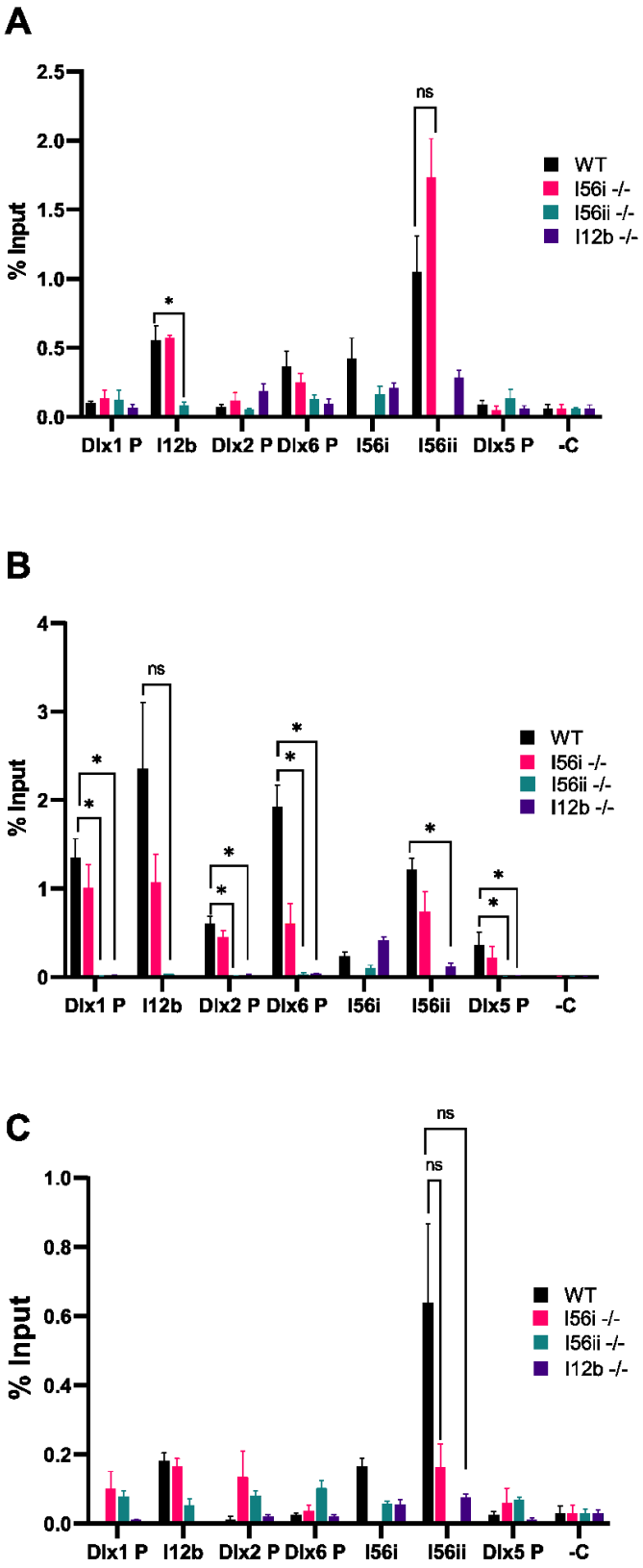
When examining H3K27Ac levels, we found that, unlike at E11.5, E13.5 embryos show a significant decrease in H3K27Ac enrichment (Fig. 4.6A) of about 75 ($p = 0.036$) exclusively at the I12b enhancer in ΔI56ii mice. These changes do track with the observed decrease in *Dlx1* expression in these mice Curiously levels also go up by about 50% (ns) at I56ii in ΔI56i mice. This is one of the few examples of an increase in enrichment that we observed between WT and mutant mice. These data indicate that the I56i and I56ii enhancer may play a role in shaping enhancer acetylation at E13.5

When examining H3K4me3 enrichment, a decrease in enrichment at *Dlx* promoters is still observed in ΔI56ii mice, as it was at E13.5. At this time point however, the decrease in H3K4me3 now is also found at the I12b and I56i enhancer. (Figure 4.6 B). It is interesting that the changed in H3K4me3 were experienced in the promoter regions first, and then the enhancers. This dynamic could have something to do with the mechanism by which demethylation is occurring at enhancer versus promoter regions. At E13.5, these broad decreases in H3K4me3 are now also observed in the ΔI12b mice. Changes in H3K4me3 are extremely severe approaching upwards of a 95% loss. In these areas enrichment reaches negligible levels (lower than the -c region). These results suggest that the timing of changes to the chromatin are specific to the genotype. Indeed ΔI12b mice showed no changes in any of the

studied modified histones at E11.5 but ΔI56i and ΔI56ii mice did. In the ΔI56i mice, H3K4me3 remains relatively comparable to WT levels. ΔI56i mice thus experience an increase in H3K27Ac at I56ii, and no changes to their H3K4me3.

Lastly, we examined H3K4me1 in all 3 genotypes. No significant changes in H3K4me1 enrichment are observed in any of the genotypes at E13.5; save for one nonsignificant drop in enrichment seen at I56ii in ΔI12b and ΔI156i mice. The large error bar associated with this data point suggests that this observation may be artefactual.

**Figure 4.6** Histone enrichment at E13.5 of WT versus *Dlx* CRE mutant mice**.** ChIP-qPCR of E13.5 ventral telencephalon for H3K27Ac A) and H3K4me3 B), and H3K4me1 C) on WT (black), ΔI56i (pink) ΔI56ii (green) and ΔI12b (purple) mice. Data are presented as %Input and normalised to the negative control region (Hoxc10I2). Data are presented as mean ± SE, n ≥ 2. Statistical analysis was done via a two tailed t-test on normalized %Input values using Holm-Sidak multiple comparisons method. * = $p \leq 0.05$


A
% Input
2.5
2.0
1.5
1.0
0.5
0.0
ns
*
WT
I56i -/-
I56ii -/-
I12b -/-
Dlx1 P
I12b
Dlx2 P
Dlx6 P
I56i
I56ii
Dlx5 P
-C
B
% Input
4
3
2
1
0
ns
*
WT
I56i -/-
I56ii -/-
I12b -/-
Dlx1 P
I12b
Dlx2 P
Dlx6 P
I56i
I56ii
Dlx5 P
-C
C
% Input
1.0
0.8
0.6
0.4
0.2
0.0
ns
ns
WT
I56i -/-
I56ii -/-
I12b -/-
Dlx1 P
I12b
Dlx2 P
Dlx6 P
I56i
I56ii
Dlx5 P
-C

Finally, we observed enrichment of all 4 major histone marks at E14.5 in mouse GEs. We looked at H3K27me3 levels in these mice due to the observed increase in enrichment in the WT mice at this period.

When examining H3K27Ac (Fig. 4.7A) we found, unsurprisingly, that levels were relatively stable between the genotypes as was observed at E11.5 and E13.5. The exception to this trend is a decrease in the mark at *Dlx6* in ΔI56i and ΔI56ii mice who do exhibit a significant downregulation of *Dlx6* expression at this time. Once again, these changes in enrichment do not perfectly correlate to the observable changes in gene expression because H3K27Ac doesn't change at *Dlx5* which is downregulated. H3K27Ac maintenance at the bigene cluster seems to be one of the more interestingly regulated marks in so far as how genotype specific and location specific the effects are. It also experiences both upregulations in some areas and downregulations in other depending on the genotype suggesting a particularly complex relationship between *Dlx* genes and the genes involved in H3K27 acetylation.

We then examined H3K4me3 enrichment in ΔI12b mice (Fig. 4.7B) to observe if the decreases in enrichment across the locus in this genotype that we observed at E13.5, but not E11.5 had been further altered. We found that the enrichment of this mark had largely returned to WT levels at E14.5 at all the assayed areas suggesting that the I12b enhancer plays an active role in shaping H3K4me3 chromatin at other *Dlx* regions specifically around E13.5.

Once again, H3K4me1 enrichment (Fig. 4.7C) was relatively stable at E14.5 between genotypes. We observed a 3-fold increase in enrichment ($p = 0.015$) at I56i in ΔI56ii mice but no changes at the I56ii enhancer as observed in E13.5 in ΔI56ii and ΔI12b mice. At this stage,

H3K4me1 enrichment is extremely low and does not meaningfully overcome the background signal in all assayed regions.

Lastly, H3K27me3 enrichment (Fig. 4.7D) which shows a spike in E14.5 WT embryos, is significantly downregulated at all enhancer and promoter elements of both the D*lx5/6* and *Dlx1/2* clusters in ΔI56i, and ΔI12b, but not in ΔI56ii mice. Enhancer I56i and I12b are hence both crucial in the upregulation of H3K27me3 in WT mice at E14.5. A summary of how each CRE influences chromatin is presented in Table 4.1

**Figure 4.7** Histone enrichment at E14.5 of WT versus *Dlx* CRE mutant mice**.** ChIP-qPCR of E14.5 ventral telencephalon for H3K27Ac A) and H3K4me3 B), and H3K4me1 C) and H3K27me3 C) on WT (black), ΔI56i (pink) ΔI56ii (green) and ΔI12b (purple) mice for. Data are presented as %Input and normalised to the negative control region (Hoxc10I2 for A, B, C, β-Actin for D). Data are presented as mean ± SE, n ≥ 2. Statistical analysis was done via a two tailed t-test on normalized %Input values using Holm-Sidak multiple comparisons method. * = $p \leq 0.05$

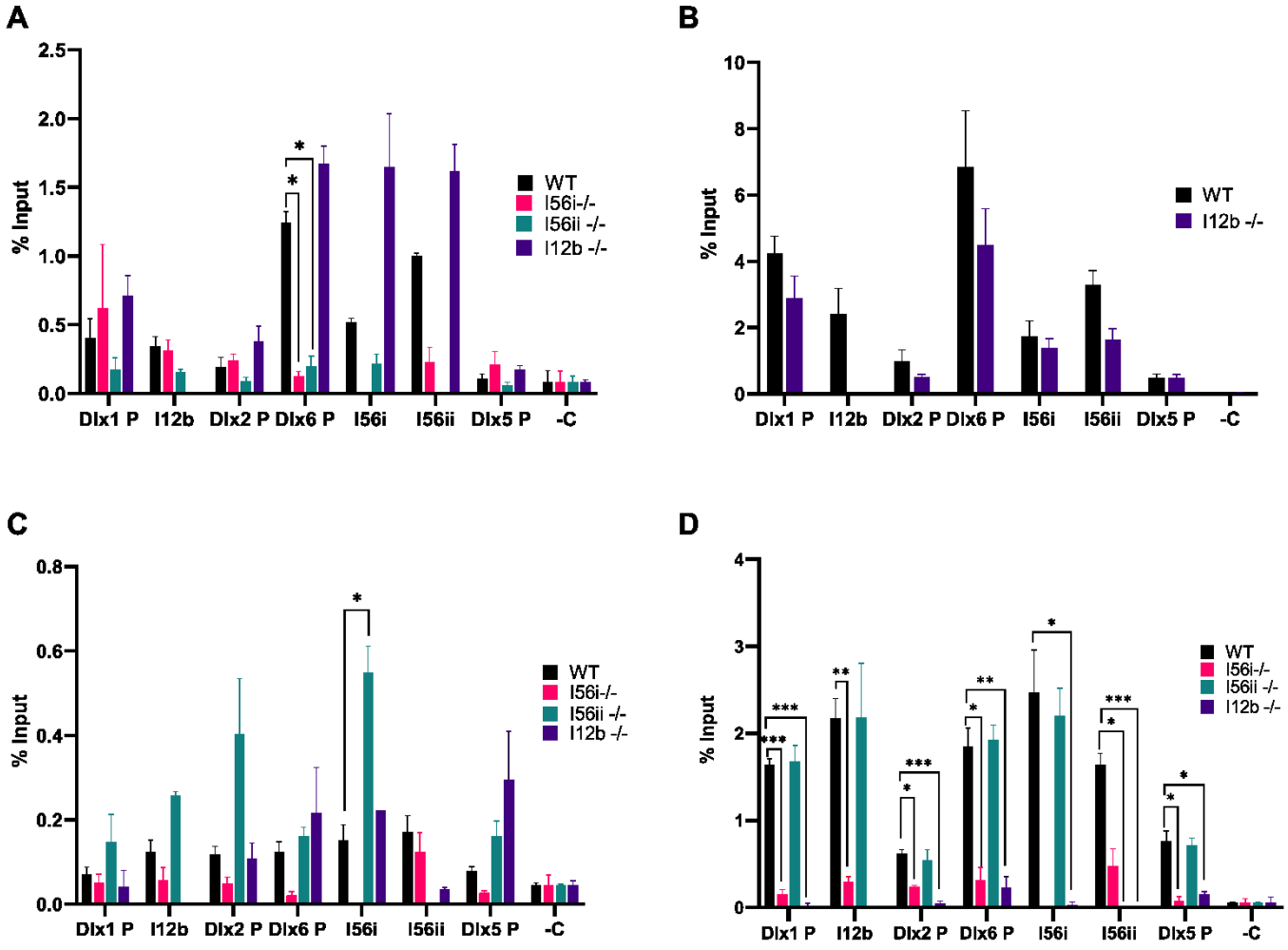

A
% Input
2.5
2.0
1.5
1.0
0.5
0.0
WT
I56i-/-
I56ii -/-
I12b -/-
Dlx1 P
I12b
Dlx2 P
Dlx6 P
I56i
I56ii
Dlx5 P
-C
B
% Input
10
8
6
4
2
0
WT
I12b -/-
Dlx1 P
I12b
Dlx2 P
Dlx6 P
I56i
I56ii
Dlx5 P
-C
C
% Input
0.8
0.6
0.4
0.2
0.0
WT
I56i-/-
I56ii -/-
I12b -/-
Dlx1 P
I12b
Dlx2 P
Dlx6 P
I56i
I56ii
Dlx5 P
-C
D
% Input
4
3
2
1
0
WT
I56i-/-
I56ii -/-
I12b -/-
Dlx1 P
I12b
Dlx2 P
Dlx6 P
I56i
I56ii
Dlx5 P
-C

**Table 4.1** Summary of how chromatin changes in each *Dlx* CRE mutant**.** How chromatin modification levels are changed in ganglionic eminences corresponding to mice of each given genotype at A) E11.5 GEs, B) E13.5 GEs, C) E14.5 relative to WT levels.

**A**

| Genotype | H3K27Ac | H3K4me3 |
|---|---|---|
| **ΔI56i** | No change | Increases at promoters of *Dlx1*, *Dlx2*, *Dlx6* |
| **ΔI56ii** | No change | Decreases at promoters of *Dlx1*, *Dlx2*, *Dlx5*, *Dlx6* |
| **ΔI12b** | No change | No change |

**B**

| Genotype | H3K27Ac | H3K4me3 | H3K4me1 |
|---|---|---|---|
| **ΔI56i** | Increases at enhancer I56ii | No Change | Decreases at enhancer I56ii |
| **ΔI56ii** | Decreases at enhancer I12b | Significant decreases everywhere | No change |
| **ΔI12b** | No change | Significant decreases everywhere | Decrease at enhancer I56ii |

**C**

| Genotype | H3K27Ac | H3K4me3 | H3K4me1 | H3K27me3 |
|---|---|---|---|---|
| **ΔI56i** | Decreases at promoter of *Dlx6* | N/A | No change | Significant decreases everywhere |
| **ΔI56ii** | Decreases at promoter of *Dlx6*, minorly at enhancer I12b | N/A | Increases at enhancer I56i | No change |
| **ΔI12b** | No change | No change | No change | Significant decreases everywhere |

**4.5 Histone modifier gene expression changes, but not dramatically, in the *Dlx* CRE mutants suggesting the CREs play a more significant role in recruiting the modifying proteins to their target histones.**

We wondered if the genotype-specific changes in methylation and acetylation levels we were observing were due to genotype specific changes in expression of the histone modifiers or if the deletion of the CREs was responsible for driving changes in the epigenetic landscapes via differential recruitment of the modulators to the chromatin. To address this question, we preformed qRT-PCR on the GEs of mice of all genotypes between E11.5-E14.5 and compared them to their WT littermates.

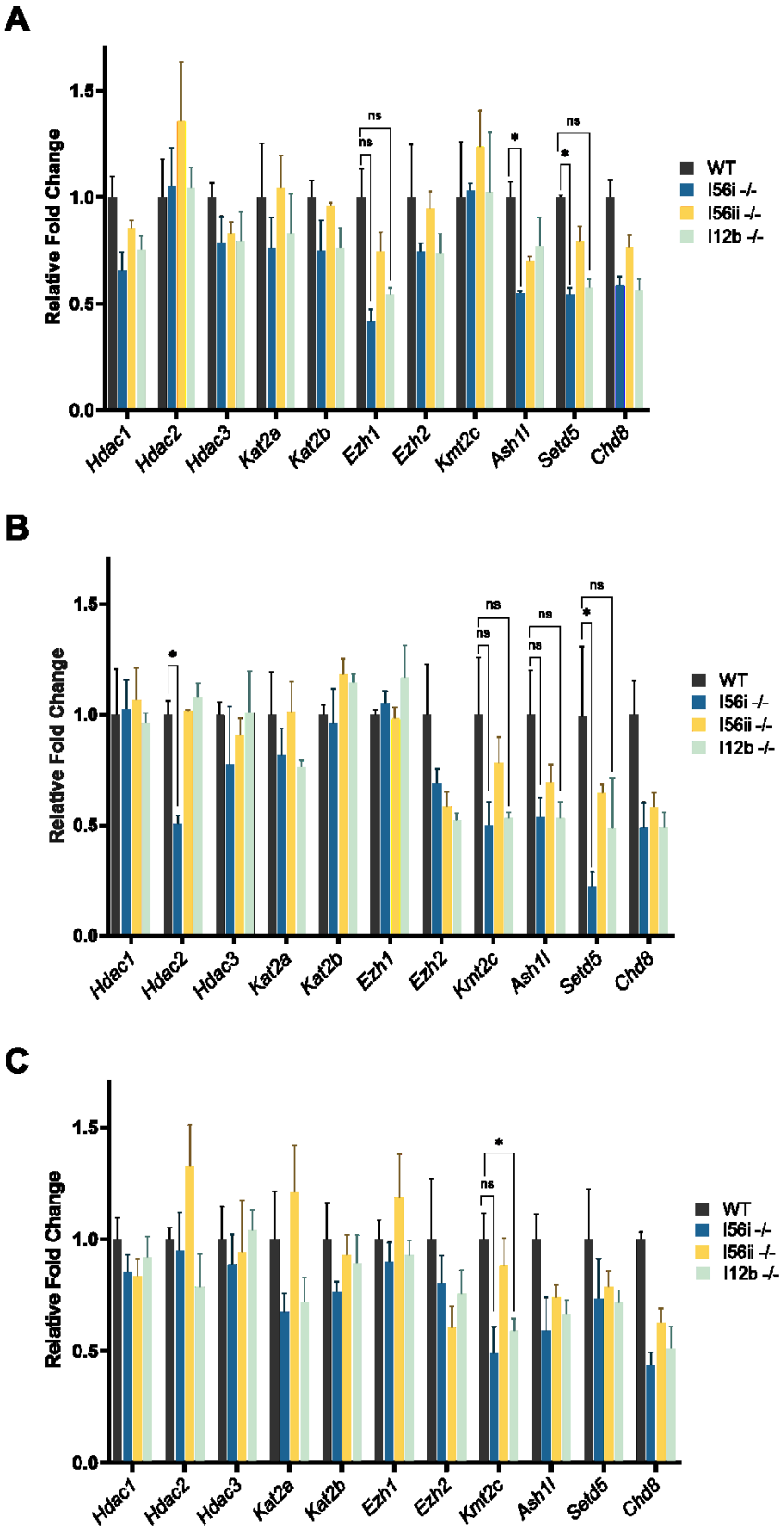
We observed a few notable changes in expression at each time point. Firstly, at E11.5 (Fig. 4.8A), we noticed a down regulation of the Polycomb group protein *Ezh1* in ΔI56i and ΔI12b mice. We did not assay mutant embryos at E11.5 for H3K27me3 but it is unlikely either of these mice experience a decrease in signal at E11.5 given that they both experience a significant upregulation of the signal at E14.5 and that the signal at E11.5 is already incredibly low. Preliminary findings (data not shown) suggest that both of these models have H3K27me3 levels that are roughly equivalent to WT levels. We also notice a decrease in expression of methyltransferases *Ash1l* ($p = 0.011$) and *Setd5* ($p = 0.026$) by about 40% in ΔI56i mice; as well as a nonsignificant reduction ($p = 0.067$) of about 35% in *Setd5* expression in ΔI12b mice.

At E13.5 (Fig. 4.8B) expression of *Ash1l* and *Setd5* is still lower than WT levels in ΔI56i and ΔI12b mice, but the difference only reaches significance for *Setd5* in ΔI56i mice ($p = 0.031$). *Ezh1* expression is back to WT levels in these mice. Expression of methyltransferase *Kmt2c* is

also reduced by about 50% in ΔI56i and ΔI12b mice at this time but changes fail to reach significance. Interestingly, *Hdac2* expression is also attenuated by about 55% in ΔI56i mice ($p = 0.031$). These changes do correspond to upregulation of H3K27Ac in these embryos at the I56ii enhancer.

By E14.5 (Figure 4.8C), expression of *Ash1I* and *Setd5* in ΔI56i and ΔI12b mice is beginning to return closer to WT levels; while *Kmt2c* transcripts remain downregulated in ΔI56i ($p = 0.078$) and ΔI12b ($p = 0.033$) embryos Expression of *Hdac2* in ΔI56i mice has returned to WT levels. H3K27me3 enrichment is downregulated in these mice but neither *Ezh1* or *Ezh2* expression is affected. In total, the majority of the changes observed in expression do not directly correlate with the changes in expression to the modifying enzymes; thus it is more likely that the CREs play a role in directing the already present modifiers to their proper targets. This is reinforced by the changes in modified histones being highly dependent on the location they are associated with in many cases instead of always being changed across the whole loci.

**Figure 4.8.** Changes in Histone modifier gene expression in *Dlx* CRE mutants**.** qRT-PCR performed on E11.5 A), E13.5 B) and E14.5 C) ganglionic eminences. Assays were run on WT (black), ΔI56i (blue), ΔI56ii (yellow), ΔI12b (green) mice. Expression values are given as fold change relative to the WT expression for each genotype and presented as mean ± SD for an n of 3. $\Delta\Delta C_q$ was calculated using the Pfafll method, normalizing to the geometric mean of *β-Actin* and *GapdH* expression. Statistical analysis was done via a two tailed t-test on ΔCq values using Holm-Sidak multiple comparisons method. * = $p \leq 0.05$


A
B
C
Relative Fold Change
1.5
1.0
0.5
0.0
ns
*
WT
I56i -/-
I56ii -/-
I12b -/-
Hdac1
Hdac2
Hdac3
Kat2a
Kat2b
Ezh1
Ezh2
Kmt2c
Ash1l
Setd5
Chd8

**4.6 Another histone modification, H3K9Ac, is affected in ΔI56i mice**

We noted a 50% increase in H3K27Ac enrichment at enhancer I56ii in E13.5 ΔI56i mice (Fig 4.6A). While this increase in enrichment did not reach significance, it did stand out this was the only increase too acetylation we observed. The increase in H3K27Ac enrichment correlates with the specific downregulation of *Hdac2* and *Setd5* which is found only in this genotype and at that point in development. We wondered if another histone acetylation mark might also be affected. We choose H3K9Ac as it is one of the most abundant acetylation marks next too H3K27Ac and it is also found at both active enhancers and active promoters (Karmodiya et al., 2012; Gates et al., 2017).

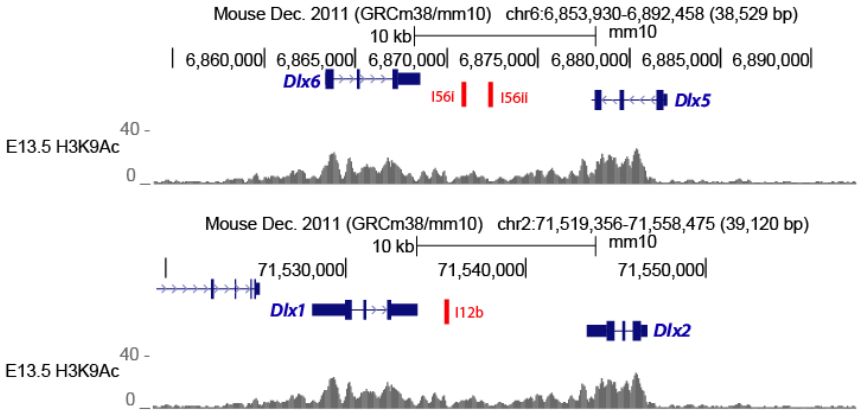
We first used *in silico* analyses to find out if the abundance of H3K9Ac in the WT mouse forebrain was similar to that of H3K27Ac (Fig 4.9A) using the Ren Lab ENCODE database as before, We found that H3K9Ac was enriched in both *Dlx* bigene clusters with peaks corresponding to the promoter and intergenic regions. Specifically, H3K9Ac was less abundant than H3K4me3 and H3K27Ac but more abundant than H3K27me3 and H3K4me1 in the forebrain. We noticed that it had strong peaks across midgestation and in the postnatal forebrain. The distribution of the peaks was similar in shape to H3K27Ac with a relatively low basal background signal in regions that did not correspond to intergenic enhancers or gene elements.

We then performed ChIP-qPCR on E13.5 GEs in WT and ΔI56i mice (Fig 4.9B). We found that H3K9Ac enrichment was not very abundant in the GEs with a very low detectable signal at all points we tested in the WT. More interestingly however, it was significantly upregulated in

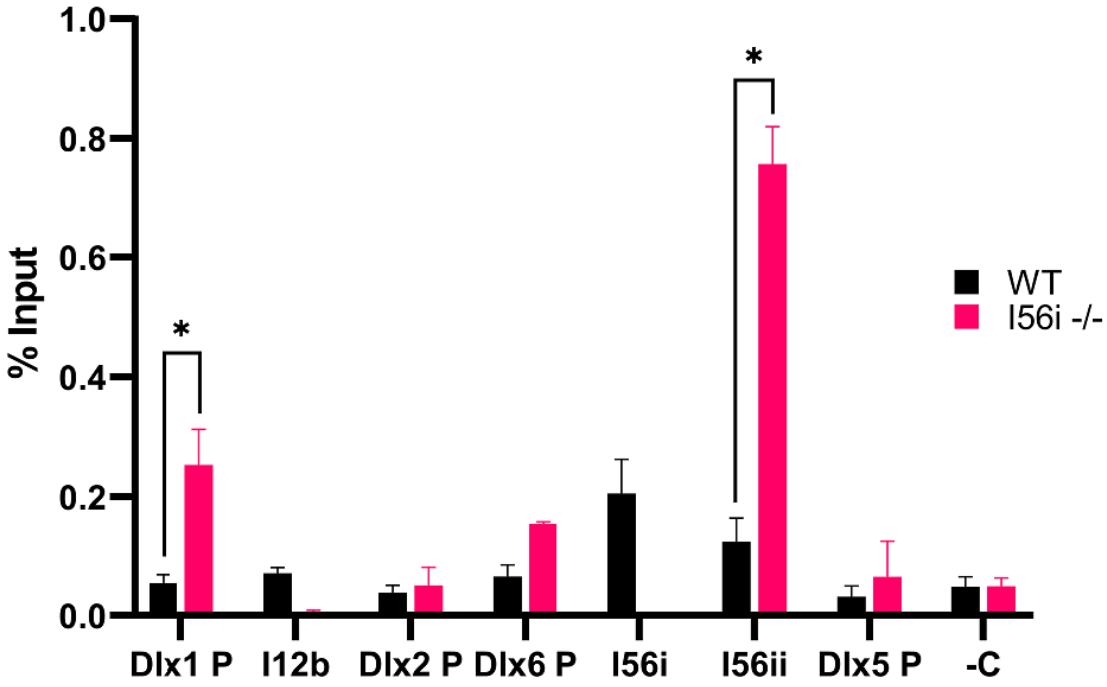
ΔI56i embryos; reaching 500% increases in enrichment at I56ii and 300% increases at the *Dlx1* promoter in the ΔI56i mice. These increases in H3K9 acetylation, especially those at the I56ii enhancer mirror the increases we observed in H3K27Ac in E13.5 GEs. Together, all these data for H3K acetylation indicate the I56i CRE is responsible for shaping the acetylation mark at more than one residue at areas corresponding to both *Dlx* clusters. We wonder if the acetylation at other loci may also be affected by these changes.

**Figure 4.9** H3K9Ac enrichment in ΔI56i mice. *In silico* data A) were acquired from the ENCODE database and visualised on the UCSC genome browser. Data represent published results from the Ren Lab (2016) mapped to the most recent mouse genome release (mm10). The data represent at least a 2X coverage. Enrichment level scale is labelled on the Y axis of each plot. Data presented as average enrichment without SD (whiskers) for clarity. The location of the Intergenic enhancers was annotated manually. ChIP-qPCR data b) are presented as %input and normalised to the negative control region (Hoxc10I2) performed on the GEs of WT (black) and ΔI56i (pink) mice at E13.5. Data are presented as mean ± SE, n = 3. Statistical analysis was done via a two tailed t-test on normalized %Input values using Holm-Sidak multiple comparisons method. * = $p \leq 0.05$

A


Mouse Dec. 2011 (GRCm38/mm10) chr6:6,853,930-6,892,458 (38,529 bp)
10 kb
mm10
6,860,000
6,865,000
6,870,000
6,875,000
6,880,000
6,885,000
6,890,000
Dlx6
I56i
I56ii
Dlx5
E13.5 H3K9Ac
40
0
Mouse Dec. 2011 (GRCm38/mm10) chr2:71,519,356-71,558,475 (39,120 bp)
10 kb
mm10
71,530,000
71,540,000
71,550,000
Dlx1
I12b
Dlx2
E13.5 H3K9Ac
40
0

B


% Input
1.0
0.8
0.6
0.4
0.2
0.0
*
*
WT
I56i -/-
Dlx1 P
I12b
Dlx2 P
Dlx6 P
I56i
I56ii
Dlx5 P
-C

## 5. Discussion

### 5.1 *Dlx* genes of mice have a stable chromatin landscape in the forebrain that doesn't significantly change during midgestation to birth.

Regulation of the *Dlx* genes has been well studied from the perspective of traditional transcriptional activation. A lot of work has been done to characterise genes which interact with the *Dlx* promoters as well as the intergenic CREs that drive expression of these genes specifically in the forebrain. However, my study is the first one to examine regulation at the level of the chromatin *in vivo*. Some of the key findings of my study are expanded upon below.

In order to understand how the chromatin might be shaped during development, we took advantage of the ENCODE database which has reported an enormous amount of ChIP-Seq data from various model organisms for most of the significant histone modifications. These data sets are not; however, without their own limitations. Firstly, some samples and libraries generated from the ENCODE project do not meet the most stringent checkmarks established by ENCODE to ensure high quality data control. Some of the issues with the ENCODE data that I reviewed in my thesis include: poor library construction (measured as the number of reads after duplicates are removed divided by the total number of reads), and severe bottlenecking (a measure of the number of loci with exactly one read relative to the number of loci with multiple reads). Wherever possible I used the best data sets that were available to me, but I did encounter these issues in the accessed ENCODE libraries for some histone marks, especially in the E11.5 forebrain. Secondly, the ENCODE project only looks at larger heterogenous tissue samples where a high degree of masking effects may be occurring. For instance, heterogeneity

may lead to confusion of regional repression of H3K27me3 versus activation through H3K27Ac. We were forced to analyse ENCODE data obtained from whole forebrain samples as this was the most specific sample set available to us.

Nonetheless the *In-silico* data do present a meaningful picture of chromatin complexity. The most obvious result is that H3K4me3 enrichment predominates throughout both *Dlx1/2* and *Dlx5/6* loci over H3K4me1, (Fig 4.1C, D, E, and F) including at the intergenic enhancers which we might expect to be H3K4me1 marked. The fact that lncRNA *Evf2* promoter is in the same region as I56i enhancer, and its gene regions extends upstream of *Dlx6* may also play a role in the prevalence of H3K4me3 over H3K4me1 in the intergenic region of the *Dlx5/6* cluster. However, no such parallel situation exists within the *Dlx1/2* cluster.

One explanation for this predominance of H3K4me3 over H3K4me1 may have to do with the 3D complexity of the chromatin in both clusters. As I mentioned in the introduction, it is not uncommon to find H3K4me3 marked enhancers when a great deal of RNA Pol II occupancy is occurring. Perhaps the consistent activation of these enhancers in cis, and potentially in trans, keep these intergenic regions locked in a chromatin configuration during development that keeps K4 trimethylated instead of monomethylated (Koch & Andrau, 2011). It could also be explained by a potential other feature of these enhancers. I mentioned earlier how CpG methylation states can be important regulators for the proper recruitment or repellant of the different histone methyltransferases. One wonders if the CpG methylation levels of this enhancer may be closer in resemblance to those of a typical promoter, and this could initiate a cascade that ultimately leads to a preference for trimethylated K4 instead of mono-methylated K4.

Next, we find that moderate levels of H3K27me3 is concomitant with H3K27Ac (Fig. 4.1A, B, G, and H) which suggests regional masking effects in the forebrain since these modifications occur on the same residue (K27) and are thus mutually exclusive (Reynolds et al, 2012; King et al., 2016). These regional masking effects may be as simple as regions where *Dlx* is actively repressed with H3K27me3, and regions where it is actively expressed with associated H3K27Ac chromatin. However, the establishment of bivalent chromatin observed in our own ChIP-qPCR data does cast some aspersion on a simple H3K27me3-H3K27Ac switch.

It is also interesting to note that there is no clear discrimination between the regions of chromatin corresponding to active enhancers, and those which are not active in this tissue. Take for example the *Dlx1/2* intergenic region which contains the I12b forebrain enhancer, and I12a which does not exhibit forebrain activity; but rather is active in the branchial arch mesenchyme (Park et al., 2004). In this case, for each histone modification examined, I12b and I12a were similarly enriched. While H3K4me1 is known to mark enhancers that may not be active, it is H3K4me3 that predominates in this cluster, and when examining H3K4me3 we still see no clear change in signal between the active and inactive enhancers, they both have strong peaks.

The prevailing abundance of H3K4me3 over H3K4me1 at these loci as well as a lack of clear differences between an active and an inactive enhancer in the *Dlx1/2* intergenic region both suggest that the modification to the *Dlx* promoters predominate over the modification to the *Dlx* CREs. It is possible that the short distance between these CREs and promoters makes discrimination between these regions difficult Especially if they are regularly sharing RNA Pol II occupancy.

During development, between E11.5 and P0, the *Dlx*-associated chromatin in the forebrain is relatively stable with slightly stronger H3K27me3 enrichment at E11.5 than at P0 whereas enrichment for H3K27Ac and H3K4me3 becomes slightly weaker during the same period. These results are a little surprising given that the expression domains of *Dlx* at E11.5 are a larger proportion of the total forebrain than at P0 when *Dlx1/2* and *Dlx5/6* expression is much more restricted in the postnatal and adult forebrain (Saino-Saito, Berlin & Baker, 2003; Wang et al., 2010). These changes during development; however, are relatively modest, and since levels of enrichment cannot be directly compared between different assays, these results may in fact be negligible. The take home picture of these chromatin landscapes is that they do not change significantly during development despite broad changes in *Dlx* expression levels and in a restriction of the domains of expression. Rather the chromatin appears to be locked into a specific configuration irrespective of the expression occurring.

### 5.2 The *Dlx1/2* and *Dlx5/6* genes are marked as bivalent chromatin in GEs and do not have their CREs associated with a H3K4me1 signal

To better understand chromatin landscapes in *Dlx*-expressing tissues of the forebrain, we performed our ChIP experiments on tissue isolated from the GE of embryos during midgestation. Our dissection technique is capable of reasonably separating out GEs from the rest of the forebrain; however, we are not able to isolate the *Dlx* expressing regions within the GEs themselves. For simplicity, we also did not study the LGE, MGE and CGE individually but rather studied them as a single unit. Therefore, compared to whole forebrain samples, we believe masking affects of a heterogenous populations are largely reduced in the current study, but we may be lacking a finer degree of detail for probing the domain specific activity of the

CREs as well as specific *Dlx1/2* versus *Dlx5/6* expression. We focused our attention once again to midgestation, E11.5-E14.5 when tangential migration of GABAergic progenitors is taking place, all four *Dlx* genes are highly expressed, and the CREs are displaying specific activity in the various subdomains of the GEs (Ghanem et al., 2003; Ghanem et al., 2007; Ghanem et al., 2008). The precise activity of these enhancers can be visualised in Fig 1.5

We found that, during development, H3K27Ac increases modestly between E11.5-E14.5 (Fig 4.3A), but the overall level is relatively low (about 16X lower than β-Actin enrichment at E14.5 in some cases, data not shown). Increases in enrichment were localised to the *Dlx*5*/6* bigene cluster but not the *Dlx1/2* bigene cluster in line with increasing transcription of these genes later into midgestation. H3K27me3 enrichment (Fig. 4.3D) predominates over H3K27Ac at E14.5 in both *Dlx5/6* and *Dlx1/2* clusters despite widespread expression of these genes. This expression is concomitant with high levels of H3K4me3, which also experiences a spike at E14.5.

The most interesting feature in these data is the widespread prevalence of both H3K27me3 and H3K4me3 is indicative of a bivalent promoter, especially since the H3K27me3 enrichment is weaker than the H3K4me3 enrichment indicating that the locus is kept in an active state. Traditionally, bivalent domains are predominately observed in ES cells; however recent publications have observed bivalent domains in mature and progenitor neurons von Schimmelmann et al., 2016; Södersten et al., 2018). These domains seem to correlate with moderate expression levels (less than chromatin marked by H3K4me3 alone). This contrasts with bivalent domains in ES cells which are generally found to be associated with actively repressed genes that are kept in a “poised” state so they can be rapidly activated. (Bernstein et al., 2006). We cannot preclude the possibility that H3K27me3 enrichment is not maintained at a

basal level in the subsections of the GEs that do not express *Dlx* genes, whereas H3K4me3 is enriched in the ventricular and subventricular zones where the genes are expressed.

The widespread enrichment of H3K27Ac in the whole forebrain ENCODE data compared to the relatively low enrichment in the data obtained from isolated GEs, despite *Dlx* gene expression here making up a large proportion of the tissue, is challenging to explain. It is possible that some regions of the forebrain that were not isolated during our extraction and that are expressing *Dlx* are epigenetically regulated by H3K27Ac when expressed and H3K27me3 when not expressed. The ENCODE data combined with our own data in the GEs suggests the *Dlx* genes in the forebrain could be found in two active states of *Dlx* chromatin, one which is H3K27Ac-enriched and one which is K4/K27-enriched. It is possible that *Dlx* gene regulation during tangential migration is dependent upon a bivalent chromatin system for a precise timing and regionalisation of gene expressions when these cells reach their final destination, the bivalence is lost and replaced with either H3K27Ac or H3K27me3.

Together my data suggest a model of chromatin regulation with the following features. First, the developmental chromatin state of the *Dlx1/2* and *Dlx5/6* bigene clusters in the GEs are bivalently marked during midgestation and experience increases in both H3K4me3 enrichment and H3K27me3 enrichment over time. We note that despite relatively stable ratios of H3K4me3/H3K27me marks during midgestation, the absolute levels of H3K4me3 enrichment are much stronger later in gestation. In agreement with these findings, a previous study from Lim et al., (2009) showed that *Dlx*2 chromatin in cultured medium spiny neurons, which normally occupy the striatum and express *Dlx* genes, contained both H3K27me3 and H3K4me3 marks, and that MLL1 was required for the *Dlx2* gene to have much higher levels of H3K4me3

enrichment over H3K27me3 enrichment. In this paper, loss of MLL1 function led to H3K27me3 prevailing over H3K4me3 and subsequent transcriptional silencing of *Dlx2*. Second, H3K27me3, which marks repressive chromatin, is necessarily more abundant than H3K27Ac in the ventral telencephalon, but this trend was not observed in the whole forebrain. We hence posit that bivalent chromatin may be a regulator of *Dlx* expression in only certain domains of the brain. Finally, we find that within the intergenic regions, the CREs do not exhibit high levels of H3K4me1 enrichment, but instead are broadly marked with H3K4me3. We postulate that this is a consequence of one or more of a few reasons: 1) short length of the cluster makes discrimination between elements difficult, 2) the high degree of interactions occurring between enhancers and promoters potentially sharing RNA Pol II occupancy, and 3) sequence specific features of the enhancers that lead to a preference for tri- over mono-methylation at K4.

**5.3 Expression levels of histone modifying genes remain relatively stable during midgestation in the mouse GEs.**

Now that we had observed some interesting chromatin landscapes, with concurrent H3K27me3 and H3K4me3 enrichment increases at late midgestation, we wondered if the increases in enrichment may be due to tissue-specific changes in the expression of the histone modifying proteins or if changes were a result of differential recruitment of the proteins to the loci. To tease this out, we preformed qRT-PCR on GEs of WT mice at E11.5, E13.5, and E14.5 when cortical neurogenesis is occurring (Fig. 4.4). We found a decrease in histone deacetylase *Hdac1* of about 50% between E11.5-E14.5 and a decrease in acetyltransferase *Kat2b* by a similar amount. Since these two genes play antagonistic roles (reviewed in Marmorstein and Zhou., 2014) we believe that any consequences of changes in their expression may somewhat

cancel each other out. However, the basal level of *Hdac1* is much higher than *Kat2b* thus a much larger degree of deacetylation activity is lost. These changes which occur between E11.5 and E13.5 do not correspond with increases in H3K27Ac at E13.5 but do not correspond with increases at E14.5. Perhaps this increase in K27 acetylation at E14.5 is a delayed response to the changes in *Hdac2* expression. On the other hand, H3K27me3 and H3K4me3 increases do not correspond with any changes in methyltransferase activity we assayed and thus are likely recruitment specific events.

### 5.4 During development the CREs play an important role in providing a stable epigenetic profile to the *Dlx* bigene clusters in the GEs

We wished to understand if the *Dlx* CREs that are active in the forebrain provided a role in *Dlx* expression that went beyond the direct activation of *Dlx* transcription and conferred epigenetic control by establishing a specific histone profile. In recent years, a great deal of emphasis has been given to how gene deletions can rewrite the topology of distal enhancers, but not on how deleting enhancers can affect the chromatin of the genes they activate. A recent paper by Huang and colleagues (2018) examined super-enhancers and found that deletions of fragments of these enhancers could have significant effects on the local chromatin via mechanisms yet to be understood. In some sense, the *Dlx* clusters make a great case study because they are relatively short and well characterised. To do this, we preformed ChIP-qPCR experiments on mice which were missing one of three intergenic enhancers: I56i, I56ii or I12b. We then preformed assays for H3K27Ac, H3K4me3, H3K4me1, and H3K27me3 on the GEs of these mice during midgestation. For each time point, we focused our efforts on the histones whose enrichment was more abundant and appreciably higher than background levels. For this

reason, we ignored H3K4me1 enrichment at E11.5, and H3K27me3 enrichment at E11.5 and E13.5.

We discovered that the CREs play significant and differential roles during development in establishing the proper epigenetic profiles. When I56i was absent, we found that mice had modest increases in H3K27Ac enrichment (Fig. 4.6A) and H3K9Ac enrichment (Fig. 4.9B) at the I56ii enhancer and at the *Dlx1* promoter specifically at E13.5, but not at earlier or later stages. Concurrent with these findings, we found a 55% decrease in *Hdac2* deacetylase expression (Fig. 4.8 B). These results suggest that the I56i enhancer is associated with *Hdac2* expression in some manner. It is possible that in the tissues I56i drives *Dlx* expression one of the *Dlx* genes is uniquely able to activate *Hdac2*. It is also possible that this decrease in deacetylation activity is required as a consequence of a compensation effect due to decreases in acetylation associated with the loss of *Evf-2* (Cajigas et al., 2015) which we have observed in these mice (Esau, 2013). By E14.5 these effects are attenuated, suggesting some special interaction occurring at this time point.

Following deletion of I56ii or of I12b, mice displayed little changes to their H3K27Ac enrichment during development. However, both mice experienced significant losses to their H3K4me3 (Fig. 4.5B and 4.6B), and these occurred earlier in the ΔI56ii mice (E11.5) than in the ΔI12b mice (E13.5), even though *Dlx1/2* are expressed before *Dlx5/6*. In ΔI12b mice, the changes in enrichment are observed only at E13.5 before returning to normal at E14.5. Unfortunately, we did not assay for H3K4me3 in ΔI56ii mice and thus cannot confirm if this chromatin dynamic is developmentally restricted to E13.5 in the ΔI56ii mutants. Nonetheless, these results, together with the changes in ΔI56i enrichment at E13.5, suggest that the

forebrain *Dlx* CREs are especially important in shaping the permissive chromatin profiles of both loci at E13.5 compared to E11.5 and E14.5. It is possible that E13.5 is a significant developmental time point for the transition from a predominant expression of *Dlx1/2* towards a predominate expression of *Dlx5/6* and marked by a unique interaction of the *Dlx1* locus with the D*lx5/6* locus at this time. At this time, it remains to be determined if the permissive chromatin marks returning-back to WT levels at E14.5 in is a function of a compensatory mechanisms, or a time point specific event where the chromatin is highly plastic during midgestation.

H3K27me3 decreases significantly at E14.5 in ΔI56i and ΔI12b mice but not in ΔI56ii mice (Fig. 4.7D). Like H3K4me3, H3K27me3 increases at E14.5 in WT mice across both loci, but in these mutants, the magnitude of the increase is marginal. In this sense, the *Dlx* genes of the ΔI56i and ΔI12b mutants essentially lose their bivalent nature and become predominately H3K4me3-enriched. Expression of the *Dlx* genes in the mutant mice, however, remains impaired at E14.5 as it is in E11.5 (Esau, 2013). Therefore, the chromatin changes we observed do not seem to result with any rescue of the expression phenotype. These data seem to reinforce the importance of bivalent chromatin in *Dlx* expression in the forebrain.

Consequently, the maintenance of bivalence during midgestation seems like an important trigger to fine tune *Dlx* expression and the CREs seem to play a significant role in establishing this bivalence, both through permissive (H3K4me3) and repressive (H3K27me3) chromatin components. We note that, with respect to bivalent chromatin, ΔI12b mice experience changes in H3K4me3 enrichment at E13.5 and H3K27me3 enrichment at E14.5; whereas ΔI56ii may be involved only in H3K4me3 enrichment and ΔI56ii involved only in

H3K27me3 enrichment. Interestingly, but perhaps unsurprisingly, ΔI12b mice which experience concurrent decreases in both H3K27me3 and H3K4me3 enrichment have the mildest phenotype of all the mutant lines.

**5.5 The CREs do not play a significant role in the expression of histone modifiers**

As mentioned previously, ΔI56i mice exhibit a decrease in *Hdac2* expression at E13.5 along with modest increases of H3K9Ac and H3K27ac enrichment (Figure 4.8 B). ΔI56i mice also show a downregulation of *Setd5*, a methyltransferase that, when downregulated, is often associated with increased acetylation levels (Osipovich, 2018). These results indicate that the I56i CRE does play some role in *Hdac2* expression and consequently in the levels of H3K acetylation across *Dlx* loci and possibly at other targets.

However, outside of this example, there are relatively little changes in the expression of the modifying histones between the genotypes. Moderate decreases in several histone methyltransferases are observed in the ΔI56i and ΔI12b mice during midgestation but these differences never exceed 2-fold. Collectively my data indicate that changes in the chromatin in the different genotypes are a consequence of the deletion of the CREs themselves, and not of the expression levels of the modifying proteins due to changes in *Dlx* expression or expression of *Dlx* targets. It is presently unclear how CREs could be responsible for affecting the recruitment of histone modifying proteins, but examining regional differences in *Dlx* targets or *Dlx* promoter interactions could illuminate this. Our study provides greater insight into how chromatin topology can change significantly in the absence of CREs and thus how CREs are important towards stabilising the chromatin of both flanking genes and of other genes, in trans.

**5.6 Future directions**

This study provides a small glimpse into the role enhancers play in stabilising the chromatin of the genes they interact with. It will be important to uncover the mechanisms by which enhancers play this role. Largely, this would involve determining which modifying proteins are differentially recruited to the *Dlx* enhancers and promoters. Then, we would need to uncover how loss of enhancers affect these modifiers. Is the change in histones a direct consequence of changes in genes that are downstream of *Dlx* and may have chromatin modifying activity? or is it perhaps a by-product of changing interactions between the enhancer and promoter elements? To answer the former hypothesis, ChIP-qPCR experiments should be used to narrow down the modifiers, followed by co-immunoprecipitation experiments to see if they interact directly with some of the genes whose expression is affected in mutants, namely the *Dlx* genes themselves. We would expect these interactions to be regionally specific to the loci in question and not broadly distributed across the chromatin. To test the latter hypothesis about the possibility of enhancer promoter interactions being causative of the changes in modified chromatin at the loci, one could use 3C to map how chromatin interactions between enhancers and promoters of the unaffected *Dlx* regions change when one enhancer is disrupted. Finally, since we assayed only a small collection of potential modifying enzymes, a more thorough approach taken to profile all the histone modifiers through RNA-Seq could be useful. We cannot exclude the possibility that other histone modifying proteins interact with the *Dlx* genes uniquely in different tissues.

Another interesting avenue of pursuit would be to determine if any of the enhancer mutant phenotypes could be rescued by restoring the chromatin configurations back to the WT

ones. While we would expect a great deal of changes to the *Dlx* expression in these mutant lines to be a consequence of an inability of effector proteins to bind to the CREs, this study sheds some credence to the possibility that some of the phenotype may be rescued by restoring the bivalent chromatin (with a high K4/K27 ratio) to these loci by overexpression of H3K4me3 and/or H3K27me3 modifying enzymes. For example, does *Dlx1* expression go back down to WT levels in ΔI12b mice when H3K27me3 is restored?

Lastly, it would be interesting to explore the genes that are necessary for enforcing the chromatin bivalency at the locus and ultimately in ensuring H3K4me3 predominates over H3K27me3. As previously mentioned, in cultured MSNs, MLL1 has been shown to be important at the *Dlx2* locus. It would be interesting to study mice with loss of function of other well-characterised activators of *Dlx* expression in the forebrain, such as *Mash1* and *Gtf2i,* to see how the chromatin landscapes change relative to WT mice. Along with these experiments, it would be prudent to see if any of the changes that occur in our mutant mice models apply to more distal chromatin elements. For example, is the changes observed in H3K27me3 in ΔI56i E14.5 GEs unique to the *Dlx* loci, or does it occur at other regions? The findings of this study point us in many directions moving forward and give us a small glimpse of a fine detail of epigenetic regulation occurring to the *Dlx* genes expressed in the forebrain.

## 6. Conclusion

The present study sought to understand the epigenetic profiles of the *Dlx1/2* and *Dlx5/6* bigene clusters in the ventral telencephalon of the mouse during midgestation. We also sought to understand how the forebrain enhancers that are in the intergenic regions of the two clusters, namely *Dlx1/2* enhancer I12b and *Dlx5/6* enhancers I56i and I56ii affect the establishment of this chromatin.

We found that, during development both *Dlx1/2* and *Dlx5/6* are marked by a bivalent chromatin states with enrichment of both H3K4me3 and H3K27me3. A reasonably stable K4/K27 ratio is preserved at both loci during midgestation. Consequently, H3K27Ac enrichment is low, and H3K4me1 enrichment is negligible even at the *Dlx* enhancers. The *Dlx* enhancers are important in establishing the bivalent chromatin, and deletion of enhancers ΔI56i, ΔI12b, and ΔI56ii leads to attenuations in the levels of H3K27me3, H3K4me3 or both in a genotype-specific, and time-specific manner.

The changes in histone enrichment for the bivalent chromatin marks are not explainable by changes in assayed histone modifiers, suggesting that the CREs have genome-specific activity that ultimately results in recruitment of different modifying factors to the chromatin.

We believe this study is important for two reasons: firstly, and most importantly, it helps us understand more about how the *Dlx* genes are regulated in the ganglionic eminences, and thus how GABAergic neurons in the cortex get to their fate – through epigenetic regulation and the prevalence of bivalent chromatin. And, secondly, in a broader sense, this study provides an interesting example of how chromatin is shaped by enhancers, and how a loss of an enhancer

sequence can have widespread effects on both the localised chromatin around it, and distal chromatin at other loci.

**References**


Acampora, D., Merlo, G. R., Paleari, L., Zerega, B., Postiglione, M. P., Mantero, S., … Levi, G. (1999). Craniofacial, vestibular and bone defects in mice lacking the Distal-less-related gene Dlx5. *Development*, *126*(17), 3795–3809.

Alvarez-Buylla, A., & García-Verdugo, J. M. (2002). Neurogenesis in Adult Subventricular Zone. *Journal of Neuroscience*, *22*(3), 629–634. https://doi.org/10.1523/JNEUROSCI.22-03-00629.2002

Anda, F. C. de, Madabhushi, R., Rei, D., Meng, J., Gräff, J., Durak, O., … Tsai, L.-H. (2016). Cortical neurons gradually attain a post-mitotic state. *Cell Research*, *26*(9), 1033–1047. https://doi.org/10.1038/cr.2016.76

Anderson, S. A., Eisenstat, D. D., Shi, L., & Rubenstein, J. L. (1997). Interneuron migration from basal forebrain to neocortex: Dependence on Dlx genes. *Science (New York, N.Y.)*, *278*(5337), 474–476. https://doi.org/10.1126/science.278.5337.474

Anderson, S. A., Qiu, M., Bulfone, A., Eisenstat, D. D., Meneses, J., Pedersen, R., & Rubenstein, J. L. (1997). Mutations of the homeobox genes Dlx-1 and Dlx-2 disrupt the striatal subventricular zone and differentiation of late born striatal neurons. *Neuron*, *19*(1), 27–37.

Anderson, Stewart A. (2002). Determination of Cell Fate within the Telencephalon. *Chemical Senses*, *27*(6), 573–575. https://doi.org/10.1093/chemse/27.6.573

Ansari, K. I., & Mandal, S. S. (2010). Mixed lineage leukemia: Roles in gene expression, hormone signaling and mRNA processing. *The FEBS Journal*, *277*(8), 1790–1804. https://doi.org/10.1111/j.1742-4658.2010.07606.x

Bannister, A. J., & Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Research*, *21*(3), 381–395. https://doi.org/10.1038/cr.2011.22

Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D. E., Wang, Z., … Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell*, *129*(4), 823–837. https://doi.org/10.1016/j.cell.2007.05.009

Berghoff, E. G., Clark, M. F., Chen, S., Cajigas, I., Leib, D. E., & Kohtz, J. D. (2013). Evf2 (Dlx6as) lncRNA regulates ultraconserved enhancer methylation and the differential transcriptional control of adjacent genes. *Development*, *140*(21), 4407–4416. https://doi.org/10.1242/dev.099390

Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., … Lander, E. S. (2006). A bivalent chromatin structure marks key developmental gene in embryonic stem cells. *Cell*, *125*(2), 315–326. https://doi.org/10.1016/j.cell.2006.02.041

Bogliotti, Y. S., & Ross, P. J. (2012). Mechanisms of histone H3 lysine 27 trimethylation remodeling during early mammalian development. *Epigenetics*, *7*(9), 976–981. https://doi.org/10.4161/epi.21615

Bonn, S., Zinzen, R. P., Girardot, C., Gustafson, E. H., Perez-Gonzalez, A., Delhomme, N., … Furlong, E. E. M. (2012). Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nature Genetics*, *44*(2), 148–156. https://doi.org/10.1038/ng.1064

Bulfone, A., Kim, H. J., Puelles, L., Porteus, M. H., Grippo, J. F., & Rubenstein, J. L. (1993). The mouse Dlx-2 (Tes-1) gene is expressed in spatially restricted domains of the forebrain,

face and limbs in midgestation mouse embryos. *Mechanisms of Development*, *40*(3), 129–140.

Cajigas, I., Leib, D. E., Cochrane, J., Luo, H., Swyter, K. R., Chen, S., … Kohtz, J. D. (2015). Evf2 lncRNA/BRG1/DLX1 interactions reveal RNA-dependent inhibition of chromatin remodeling. *Development (Cambridge, England)*, *142*(15), 2641–2652. https://doi.org/10.1242/dev.126318

Calo, E., & Wysocka, J. (2013). Modification of enhancer chromatin: What, how and why? *Molecular Cell*, *49*(5). https://doi.org/10.1016/j.molcel.2013.01.038

Cao, J., & Yan, Q. (2012). Histone Ubiquitination and Deubiquitination in Transcription, DNA Damage Response, and Cancer. *Frontiers in Oncology*, *2*. https://doi.org/10.3389/fonc.2012.00026

Caric, D., Gooday, D., Hill, R. E., McConnell, S. K., & Price, D. J. (1997). Determination of the migratory capacity of embryonic cortical cells lacking the transcription factor Pax-6. *Development*, *124*(24), 5087–5096.

Cheng, C., Yan, K.-K., Yip, K. Y., Rozowsky, J., Alexander, R., Shou, C., & Gerstein, M. (2011). A statistical framework for modeling gene expression using chromatin features and application to modENCODE datasets. *Genome Biology*, *12*(2), R15. https://doi.org/10.1186/gb-2011-12-2-r15

Cheung, P., & Lau, P. (2005). Epigenetic regulation by histone methylation and histone variants. *Molecular Endocrinology (Baltimore, Md.)*, *19*(3), 563–573. https://doi.org/10.1210/me.2004-0496

Cho, B., Kim, H. J., Kim, H., & Sun, W. (2011). Changes in the Histone Acetylation Patterns during the Development of the Nervous System. *Experimental Neurobiology*, *20*(2), 81–84. https://doi.org/10.5607/en.2011.20.2.81

Cobos, I., Calcagnotto, M. E., Vilaythong, A. J., Thwin, M. T., Noebels, J. L., Baraban, S. C., & Rubenstein, J. L. R. (2005). Mice lacking Dlx1 show subtype-specific loss of interneurons, reduced inhibition and epilepsy. *Nature Neuroscience*, *8*(8), 1059–1068. https://doi.org/10.1038/nn1499

Cohen, S. M., Brönner, G., Küttner, F., Jürgens, G., & Jäckle, H. (1989). Distal-less encodes a homoeodomain protein required for limb development in Drosophila. *Nature*, *338*(6214), 432–434. https://doi.org/10.1038/338432a0

Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., … Jaenisch, R. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(50), 21931–21936. https://doi.org/10.1073/pnas.1016071107

Depew, M. J., Liu, J. K., Long, J. E., Presley, R., Meneses, J. J., Pedersen, R. A., & Rubenstein, J. L. (1999). Dlx5 regulates regional development of the branchial arches and sensory capsules. *Development*, *126*(17), 3831–3846.

Dollé, P., Price, M., & Duboule, D. (1992). Expression of the murine Dlx-1 homeobox gene during facial, ocular and limb development. *Differentiation; Research in Biological Diversity*, *49*(2), 93–99.

Dong, P. D., Chu, J., & Panganiban, G. (2000). Coexpression of the homeobox genes Distal-less and homothorax determines Drosophila antennal identity. *Development*, *127*(2), 209–216.

Dong, X., & Weng, Z. (2013). The correlation between histone modifications and gene expression. *Epigenomics*, *5*(2), 113–116. https://doi.org/10.2217/epi.13.13

Duboule, D, & Dollé, P. (1989). The structural and functional organization of the murine HOX gene family resembles that of Drosophila homeotic genes. *The EMBO Journal*, *8*(5), 1497–1505.

Duboule, Denis. (2007). The rise and fall of Hox gene clusters. *Development*, *134*(14), 2549–2560. https://doi.org/10.1242/dev.001065

Dunwell, T. L., & Holland, P. W. H. (2016). Diversity of human and mouse homeobox gene expression in development and adult tissues. *BMC Developmental Biology*, *16*(1), 40. https://doi.org/10.1186/s12861-016-0140-y

Eisenstat, D. D., Liu, J. K., Mione, M., Zhong, W., Yu, G., Anderson, S. A., … Rubenstein, J. L. (1999). DLX-1, DLX-2, and DLX-5 expression define distinct stages of basal forebrain differentiation. *The Journal of Comparative Neurology*, *414*(2), 217–237.

Ellies, D. L., Stock, D. W., Hatch, G., Giroux, G., Weiss, K. M., & Ekker, M. (1997). Relationship between the genomic organization and the overlapping embryonic expression patterns of the zebrafish dlx genes. *Genomics*, *45*(3), 580–590. https://doi.org/10.1006/geno.1997.4978

Emerald, B. S., & Roy, J. K. (1997). Homeotic transformation in Drosophila. *Nature*, *389*(6652), 684. https://doi.org/10.1038/39500

Esau, C. (2013). Functional characterization of *Dlx* enhancers in the developing mouse (M.Sc. dissertation, University of Ottawa (Canada)).

Fazel Darbandi, S (2014). Functional analysis of *Dlx* intergenic enhancers in the developing mouse forebrain. (PhD dissertation, University of Ottawa (Canada)).

Fazel Darbandi, S., Poitras, L., Monis, S., Lindtner, S., Yu, M., Hatch, G., … Ekker, M. (2016). Functional consequences of I56ii Dlx enhancer deletion in the developing mouse forebrain. *Developmental Biology*. https://doi.org/10.1016/j.ydbio.2016.10.015

Feng, J., Bi, C., Clark, B., Mady, R., Shah, P., & Kohtz, J. (2006). The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes & Development*, *20*, 1470–1484. https://doi.org/10.1101/gad.1416106

Fongang, B., Kong, F., Negi, S., Braun, W., & Kudlicki, A. (2016). A Conserved Structural Signature of the Homeobox Coding DNA in HOX genes. *Scientific Reports*, *6*, 35415. https://doi.org/10.1038/srep35415

Gates, L. A., Shi, J., Rohira, A. D., Feng, Q., Zhu, B., Bedford, M. T., … O'Malley, B. W. (2017). Acetylation on histone H3 lysine 9 mediates a switch from transcription initiation to elongation. *Journal of Biological Chemistry*, jbc.M117.802074. https://doi.org/10.1074/jbc.M117.802074

Gaunt, S. J. (1988). Mouse homeobox gene transcripts occupy different but overlapping domains in embryonic germ layers and organs: A comparison of Hox-3.1 and Hox-1.5. *Development (Cambridge, England)*, *103*(1), 135–144.

Gaunt, S. J., Miller, J. R., Powell, D. J., & Duboule, D. (1986). Homoeobox gene expression in mouse embryos varies with position by the primitive streak stage. *Nature*, *324*(6098), 662–664. https://doi.org/10.1038/324662a0

Gaunt, Stephen J. (2015). The significance of Hox gene collinearity. *The International Journal of Developmental Biology*, *59*(4–6), 159–170. https://doi.org/10.1387/ijdb.150223sg

Gehring, W. J., Affolter, M., & Bürglin, T. (1994). Homeodomain proteins. *Annual Review of Biochemistry*, *63*, 487–526. https://doi.org/10.1146/annurev.bi.63.070194.002415

Ghanem, N., Jarinova, O., Amores, A., Long, Q., Hatch, G., Park, B. K., … Ekker, M. (2003). Regulatory Roles of Conserved Intergenic Domains in Vertebrate Dlx Bigene Clusters. *Genome Research*, *13*(4), 533–543. https://doi.org/10.1101/gr.716103

Ghanem, N., Yu, M., Long, J., Hatch, G., Rubenstein, J. L. R., & Ekker, M. (2007). Distinct cis-regulatory elements from the Dlx1/Dlx2 locus mark different progenitor cell populations in the ganglionic eminences and different subtypes of adult cortical interneurons. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(19), 5012–5022. https://doi.org/10.1523/JNEUROSCI.4725-06.2007

Ghanem, N., Yu, M., Poitras, L., Rubenstein, J. L. R., & Ekker, M. (2008). Characterization of a distinct subpopulation of striatal projection neurons expressing the Dlx genes in the basal ganglia through the activity of the I56ii enhancer. *Developmental Biology*, *322*(2), 415–424. https://doi.org/10.1016/j.ydbio.2008.07.029

Gilbert, S. F. (2000). Differentiation of the Neural Tube. *Developmental Biology. 6th Edition*. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK10034/

Grone, B. P., & Maruska, K. P. (2016). Three Distinct Glutamate Decarboxylase Genes in Vertebrates. *Scientific Reports*, *6*. https://doi.org/10.1038/srep30507

Haberland, M., Montgomery, R. L., & Olson, E. N. (2009). The many roles of histone deacetylases in development and physiology: Implications for disease and therapy. *Nature Reviews. Genetics*, *10*(1), 32–42. https://doi.org/10.1038/nrg2485

Hong, L., Schroth, G. P., Matthews, H. R., Yau, P., & Bradbury, E. M. (1993). Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA. *The Journal of Biological Chemistry*, *268*(1), 305–314.

Hu, Z., & Tee, W.-W. (2017). Enhancers and chromatin structures: Regulatory hubs in gene expression and diseases. *Bioscience Reports*, *37*(2). https://doi.org/10.1042/BSR20160183

Huang, J., Li, K., Cai, W., Liu, X., Zhang, Y., Orkin, S. H., … Yuan, G.-C. (2018). Dissecting super-enhancer hierarchy based on chromatin interactions. *Nature Communications*, *9*(1), 943. https://doi.org/10.1038/s41467-018-03279-9

Hyun, K., Jeon, J., Park, K., & Kim, J. (2017). Writing, erasing and reading histone lysine methylations. *Experimental & Molecular Medicine*, *49*(4), e324. https://doi.org/10.1038/emm.2017.11

Jäkel, S., & Dimou, L. (2017). Glial Cells and Their Function in the Adult Brain: A Journey through the History of Their Ablation. *Frontiers in Cellular Neuroscience*, *11*. https://doi.org/10.3389/fncel.2017.00024

Kaji, T., & Artinger, K. B. (2004). Dlx3b and dlx4b function in the development of Rohon-Beard sensory neurons and trigeminal placode in the zebrafish neurula. *Developmental Biology*, *276*(2), 523–540. https://doi.org/10.1016/j.ydbio.2004.09.020

Karmodiya, K., Krebs, A. R., Oulad-Abdelghani, M., Kimura, H., & Tora, L. (2012). H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genomics*, *13*, 424. https://doi.org/10.1186/1471-2164-13-424

Katsani, K. R., Arredondo, J. J., Kal, A. J., & Verrijzer, C. P. (2001). A homeotic mutation in the trithorax SET domain impedes histone binding. *Genes & Development*, *15*(17), 2197–2202. https://doi.org/10.1101/gad.201901

Kaufman, D. L., Houser, C. R., & Tobin, A. J. (1991). Two forms of the gamma-aminobutyric acid synthetic enzyme glutamate decarboxylase have distinct intraneuronal distributions and cofactor interactions. *Journal of Neurochemistry*, *56*(2), 720–723. https://doi.org/10.1111/j.1471-4159.1991.tb08211.x

Kiecker, C., & Lumsden, A. (2005). Compartments and their boundaries in vertebrate brain development. *Nature Reviews Neuroscience*, *6*(7), 553. https://doi.org/10.1038/nrn1702

King, A. D., Huang, K., Rubbi, L., Liu, S., Wang, C.-Y., Wang, Y., … Fan, G. (2016). Reversible Regulation of Promoter and Enhancer Histone Landscape by DNA Methylation in Mouse Embryonic Stem Cells. *Cell Reports*, *17*(1), 289–302. https://doi.org/10.1016/j.celrep.2016.08.083

Kissinger, C. R., Liu, B. S., Martin-Blanco, E., Kornberg, T. B., & Pabo, C. O. (1990). Crystal structure of an engrailed homeodomain-DNA complex at 2.8 A resolution: A framework

for understanding homeodomain-DNA interactions. *Cell*, *63*(3), 579–590. https://doi.org/10.1016/0092-8674(90)90453-I

Koch, F., & Andrau, J.-C. (2011). Initiating RNA polymerase II and TIPs as hallmarks of enhancer activity and tissue-specificity. *Transcription*, *2*(6), 263–268. https://doi.org/10.4161/trns.2.6.18747

Kurimoto, K., Yabuta, Y., Hayashi, K., Ohta, H., Kiyonari, H., Mitani, T., … Saitou, M. (2015). Quantitative Dynamics of Chromatin Remodeling during Germ Cell Specification from Mouse Embryonic Stem Cells. *Cell Stem Cell*, *16*(5), 517–532. https://doi.org/10.1016/j.stem.2015.03.002

Le, T. N., Zhou, Q.-P., Cobos, I., Zhang, S., Zagozewski, J., Japoni, S., … Eisenstat, D. D. (2017). GABAergic Interneuron Differentiation in the Basal Forebrain Is Mediated through Direct Regulation of Glutamic Acid Decarboxylase Isoforms by Dlx Homeobox Transcription Factors. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *37*(36), 8816–8829. https://doi.org/10.1523/JNEUROSCI.2125-16.2017

Lesley, S., & Rodin, I. (2011). *Psychiatry* (2nd ed.). Retrieved from https://www.sciencedirect.com/book/9780702033964/psychiatry#book-info

Levine, M., Rubin, G. M., & Tjian, R. (1984). Human DNA sequences homologous to a protein coding region conserved between homeotic genes of Drosophila. *Cell*, *38*(3), 667–673. https://doi.org/10.1016/0092-8674(84)90261-7

Lim, D. A., Huang, Y.-C., Swigut, T., Mirick, A. L., Garcia-Verdugo, J. M., Wysocka, J., … Alvarez-Buylla, A. (2009). Chromatin remodelling factor Mll1 is essential for neurogenesis from

postnatal neural stem cells. *Nature*, *458*(7237), 529–533. https://doi.org/10.1038/nature07726

Liu, J. K., Ghattas, I., Liu, S., Chen, S., & Rubenstein, J. L. (1997). Dlx genes encode DNA-binding proteins that are expressed in an overlapping and sequential pattern during basal ganglia differentiation. *Developmental Dynamics: An Official Publication of the American Association of Anatomists*, *210*(4), 498–512. https://doi.org/10.1002/(SICI)1097-0177(199712)210:4<498::AID-AJA12>3.0.CO;2-3

Liu, X., Novosedlik, N., Wang, A., Hudson, M. L., Cohen, I. L., Chudley, A. E., … Holden, J. J. A. (2009). The DLX1and DLX2 genes and susceptibility to autism spectrum disorders. *European Journal of Human Genetics*, *17*(2), 228–235. https://doi.org/10.1038/ejhg.2008.148

Lomvardas, S., & Maniatis, T. (2016). Histone and DNA Modifications as Regulators of Neuronal Development and Function. *Cold Spring Harbor Perspectives in Biology*, *8*(7). https://doi.org/10.1101/cshperspect.a024208

López-Bendito, G., Cautinat, A., Sánchez, J. A., Bielle, F., Flames, N., Garratt, A. N., … Garel, S. (2006). Tangential neuronal migration controls axon guidance: A role for neuregulin-1 in thalamocortical axon navigation. *Cell*, *125*(1), 127–142. https://doi.org/10.1016/j.cell.2006.01.042

Lupiáñez, D. G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., … Mundlos, S. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell*, *161*(5), 1012–1025. https://doi.org/10.1016/j.cell.2015.04.004

Luskin, M.B., & Coskun, V. (2002). The progenitor cells of the embryonic telencephalon and the neonatal anterior subventricular zone differentially regulate their cell cycle. *Chemical Senses*, *27*(6), 577–580.

Magklara, A., Yen, A., Colquitt, B. M., Clowney, E. J., Allen, W., Markenscoff-Papadimitriou, E., … Lomvardas, S. (2011). An epigenetic signature for monoallelic olfactory receptor expression. *Cell*, *145*(4), 555–570. https://doi.org/10.1016/j.cell.2011.03.040

Marín, O., & Rubenstein, J. L. (2001). A long, remarkable journey: Tangential migration in the telencephalon. *Nature Reviews. Neuroscience*, *2*(11), 780–790. https://doi.org/10.1038/35097509

Marín, Oscar, Plump, A. S., Flames, N., Sánchez-Camacho, C., Tessier-Lavigne, M., & Rubenstein, J. L. R. (2003). Directional guidance of interneuron migration to the cerebral cortex relies on subcortical Slit1/2-independent repulsion and cortical attraction. *Development*, *130*(9), 1889–1901. https://doi.org/10.1242/dev.00417

Marmorstein, R., & Zhou, M.-M. (2014). Writers and Readers of Histone Acetylation: Structure, Mechanism, and Inhibition. *Cold Spring Harbor Perspectives in Biology*, *6*(7), a018762. https://doi.org/10.1101/cshperspect.a018762

McGinnis, W., Garber, R. L., Wirz, J., Kuroiwa, A., & Gehring, W. J. (1984). A homologous protein-coding sequence in Drosophila homeotic genes and its conservation in other metazoans. *Cell*, *37*(2), 403–408. https://doi.org/10.1016/0092-8674(84)90370-2

McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa, A., & Gehring, W. J. (1984). A conserved DNA sequence in homoeotic genes of the Drosophila Antennapedia and bithorax complexes. *Nature*, *308*(5958), 428–433. https://doi.org/10.1038/308428a0

McGuinness, T., Porteus, M. H., Smiga, S., Bulfone, A., Kingsley, C., Qiu, M., … Rubenstein, J. L. (1996). Sequence, organization, and transcription of the Dlx-1 and Dlx-2 locus. *Genomics*, *35*(3), 473–485. https://doi.org/10.1006/geno.1996.0387

Miller, T., Krogan, N. J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., … Shilatifard, A. (2001). COMPASS: A complex of proteins associated with a trithorax-related SET domain protein. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(23), 12902–12907. https://doi.org/10.1073/pnas.231473398

Milne, T. A., Briggs, S. D., Brock, H. W., Martin, M. E., Gibbs, D., Allis, C. D., & Hess, J. L. (2002). MLL targets SET domain methyltransferase activity to Hox gene promoters. *Molecular Cell*, *10*(5), 1107–1117.

Miyoshi, G., Hjerling-Leffler, J., Karayannis, T., Sousa, V. H., Butt, S. J. B., Battiste, J., … Fishell, G. (2010). Genetic fate mapping reveals that the caudal ganglionic eminence produces a large and diverse population of superficial cortical interneurons. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*(5), 1582–1594. https://doi.org/10.1523/JNEUROSCI.4515-09.2010

Nikolopoulou, E., Galea, G. L., Rolo, A., Greene, N. D. E., & Copp, A. J. (2017). Neural tube closure: Cellular, molecular and biomechanical mechanisms. *Development (Cambridge, England)*, *144*(4), 552–566. https://doi.org/10.1242/dev.145904

Oda, H., Okamoto, I., Murphy, N., Chu, J., Price, S. M., Shen, M. M., … Reinberg, D. (2009). Monomethylation of histone H4-lysine 20 is involved in chromosome structure and

stability and is essential for mouse development. *Molecular and Cellular Biology*, *29*(8), 2278–2295. https://doi.org/10.1128/MCB.01768-08

Oka, R., Zicola, J., Weber, B., Anderson, S. N., Hodgman, C., Gent, J. I., … Stam, M. (2017). Genome-wide mapping of transcriptional enhancer candidates using DNA and chromatin features in maize. *Genome Biology*, *18*. https://doi.org/10.1186/s13059-017-1273-4

O'Rourke, N. A., Dailey, M. E., Smith, S. J., & McConnell, S. K. (1992). Diverse migratory pathways in the developing cerebral cortex. *Science (New York, N.Y.)*, *258*(5080), 299–302. https://doi.org/10.1126/science.1411527

O'Rourke, N. A., Sullivan, D. P., Kaznowski, C. E., Jacobs, A. A., & McConnell, S. K. (1995). Tangential migration of neurons in the developing cerebral cortex. *Development (Cambridge, England)*, *121*(7), 2165–2176.

Osipovich, A. B., Gangula, R., Vianna, P. G., & Magnuson, M. A. (2016). Setd5 is essential for mammalian development and the co-transcriptional regulation of histone acetylation. *Development (Cambridge, England)*, *143*(24), 4595–4607. https://doi.org/10.1242/dev.141465

Park, B. K., Sperber, S. M., Choudhury, A., Ghanem, N., Hatch, G. T., Sharpe, P. T., … Ekker, M. (2004). Intergenic enhancers with distinct activities regulate Dlx gene expression in the mesenchyme of the branchial arches. *Developmental Biology*, *268*(2), 532–545. https://doi.org/10.1016/j.ydbio.2004.01.010

Perera, M., Merlo, G. R., Verardo, S., Paleari, L., Corte, G., & Levi, G. (2004). Defective neuronogenesis in the absence of Dlx5. *Molecular and Cellular Neurosciences*, *25*(1), 153–161. https://doi.org/10.1016/j.mcn.2003.10.004

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Research*, *29*(9), e45.

Pfluger, J., & Wagner, D. (2007). Histone modifications and dynamic regulation of genome accessibility in plants. *Current Opinion in Plant Biology*, *10*(6), 645–652. https://doi.org/10.1016/j.pbi.2007.07.013

Poitras, L., Yu, M., Lesage-Pelletier, C., Macdonald, R. B., Gagné, J.-P., Hatch, G., … Ekker, M. (2010). An SNP in an ultraconserved regulatory element affects Dlx5/Dlx6 regulation in the forebrain. *Development (Cambridge, England)*, *137*(18), 3089–3097. https://doi.org/10.1242/dev.051052

Poitras, L., Ghanem, N., Hatch, G., & Ekker, M. (2007). The proneural determinant MASH1 regulates forebrain Dlx1/2 expression through the I12b intergenic enhancer. *Development*, *134*(9), 1755–1765. https://doi.org/10.1242/dev.02845

Porteus, M. H., Bulfone, A., Liu, J. K., Puelles, L., Lo, L. C., & Rubenstein, J. L. (1994). DLX-2, MASH-1, and MAP-2 expression and bromodeoxyuridine incorporation define molecularly distinct cell populations in the embryonic mouse forebrain. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *14*(11 Pt 1), 6370–6383.

Potter, G. B., Petryniak, M. A., Shevchenko, E., McKinsey, G. L., Ekker, M., & Rubenstein, J. L. R. (2009). Generation of Cre-transgenic mice using Dlx1/Dlx2 enhancers and their

characterization in GABAergic interneurons. *Molecular and Cellular Neurosciences*, *40*(2), 167–186. https://doi.org/10.1016/j.mcn.2008.10.003

Price, M., Lazzaro, D., Pohl, T., Mattei, M. G., Rüther, U., Olivo, J. C., … Di Lauro, R. (1992). Regional expression of the homeobox gene Nkx-2.2 in the developing mammalian forebrain. *Neuron*, *8*(2), 241–255.

Price, M., Lemaistre, M., Pischetola, M., Di Lauro, R., & Duboule, D. (1991). A mouse gene related to Distal-less shows a restricted expression in the developing forebrain. *Nature*, *351*(6329), 748–751. https://doi.org/10.1038/351748a0

Proudfoot, A., Axelrod, H. L., Geralt, M., Fletterick, R. J., Yumoto, F., Deacon, A. M., … Serrano, P. (2016). Dlx5 homedomain/DNA complex; Structure, binding and effect of mutations related to split-hand and foot malformation syndrome. *Journal of Molecular Biology*, *428*(6), 1130–1141. https://doi.org/10.1016/j.jmb.2016.01.023

Qian, Y. Q., Resendez-Perez, D., Gehring, W. J., & Wüthrich, K. (1994). The des(1-6)antennapedia homeodomain: Comparison of the NMR solution structure and the DNA-binding affinity with the intact Antennapedia homeodomain. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(9), 4091–4095.

Qiu, M., Bulfone, A., Martinez, S., Meneses, J. J., Shimamura, K., Pedersen, R. A., & Rubenstein, J. L. (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 & Development*, *9*(20), 2523–2538. https://doi.org/10.1101/gad.9.20.2523

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S. A., Flynn, R. A., & Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*, *470*(7333), 279–283. https://doi.org/10.1038/nature09692

Rando, O. J., & Ahmad, K. (2007). Rules and regulation in the primary structure of chromatin. *Current Opinion in Cell Biology*, *19*(3), 250–256. https://doi.org/10.1016/j.ceb.2007.04.006

Reynolds, N., Salmon-Divon, M., Dvinge, H., Hynes-Allen, A., Balasooriya, G., Leaford, D., … Hendrich, B. (2012). NuRD-mediated deacetylation of H3K27 facilitates recruitment of Polycomb Repressive Complex 2 to direct gene repression. *The EMBO Journal*, *31*(3), 593–605. https://doi.org/10.1038/emboj.2011.431

Robledo, R. F., Rajan, L., Li, X., & Lufkin, T. (2002). The Dlx5 and Dlx6 homeobox genes are essential for craniofacial, axial, and appendicular skeletal development. *Genes & Development*, *16*(9), 1089–1101. https://doi.org/10.1101/gad.988402

Rossignol, E. (2011). Genetics and Function of Neocortical GABAergic Interneurons in Neurodevelopmental Disorders [Research article]. https://doi.org/10.1155/2011/649325

Russ, B. E., Olshansky, M., Li, J., Nguyen, M. L. T., Gearing, L. J., Nguyen, T. H. O., … Turner, S. J. (2017). Regulation of H3K4me3 at Transcriptional Enhancers Characterizes Acquisition of Virus-Specific CD8+ T Cell-Lineage-Specific Function. *Cell Reports*, *21*(12), 3624–3636. https://doi.org/10.1016/j.celrep.2017.11.097

Saino-Saito, S., Berlin, R., & Baker, H. (2003). Dlx-1 and Dlx-2 expression in the adult mouse brain: Relationship to dopaminergic phenotypic regulation. *The Journal of Comparative Neurology*, *461*(1), 18–30. https://doi.org/10.1002/cne.10611

Sawicka, A., & Seiser, C. (2012). Histone H3 phosphorylation – A versatile chromatin modification for different occasions. *Biochimie*, *94*(11), 2193–2201. https://doi.org/10.1016/j.biochi.2012.04.018

Schuettengruber, B., Bourbon, H.-M., Croce, L. D., & Cavalli, G. (2017). Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell*, *171*(1), 34–57. https://doi.org/10.1016/j.cell.2017.08.002

Schuettengruber, B., & Cavalli, G. (2009). Recruitment of polycomb group complexes and their role in the dynamic regulation of cell fate choice. *Development (Cambridge, England)*, *136*(21), 3531–3542. https://doi.org/10.1242/dev.033902

Schüle, B., Li, H. H., Fisch-Kohl, C., Purmann, C., & Francke, U. (2007). DLX5 and DLX6 Expression Is Biallelic and Not Modulated by MeCP2 Deficiency. *American Journal of Human Genetics*, *81*(3), 492–506.

Scott, M. P., & Weiner, A. J. (1984). Structural relationships among genes that control development: Sequence homology between the Antennapedia, Ultrabithorax, and fushi tarazu loci of Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, *81*(13), 4115–4119. https://doi.org/10.1073/pnas.81.13.4115

Seimiya, M., Ishiguro, H., Miura, K., Watanabe, Y., & Kurosawa, Y. (1994). Homeobox-containing genes in the most primitive metazoa, the sponges. *European Journal of Biochemistry*, *221*(1), 219–225. https://doi.org/10.1111/j.1432-1033.1994.tb18732.x

Sharifi-Zarchi, A., Gerovska, D., Adachi, K., Totonchi, M., Pezeshk, H., Taft, R. J., … Araúzo-Bravo, M. J. (2017). DNA methylation regulates discrimination of enhancers from promoters through a H3K4me1-H3K4me3 seesaw mechanism. *BMC Genomics*, *18*. https://doi.org/10.1186/s12864-017-4353-7

Shilatifard, A. (2012). The COMPASS family of histone H3K4 methylases: Mechanisms of regulation in development and disease pathogenesis. *Annual Review of Biochemistry*, *81*, 65–95. https://doi.org/10.1146/annurev-biochem-051710-134100

Shogren-Knaak, M., Ishii, H., Sun, J.-M., Pazin, M. J., Davie, J. R., & Peterson, C. L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science (New York, N.Y.)*, *311*(5762), 844–847. https://doi.org/10.1126/science.1124000

Södersten, E., Toskas, K., Rraklli, V., Tiklova, K., Björklund, Å. K., Ringnér, M., … Holmberg, J. (2018). A comprehensive map coupling histone modifications with gene regulation in adult dopaminergic and serotonergic neurons. *Nature Communications*, *9*(1), 1226. https://doi.org/10.1038/s41467-018-03538-9

Stevens, B. (2003). Glia: Much more than the neuron's side-kick. *Current Biology: CB*, *13*(12), R469-472.

Stock, D. W., Ellies, D. L., Zhao, Z., Ekker, M., Ruddle, F. H., & Weiss, K. M. (1996). The evolution of the vertebrate Dlx gene family. *Proceedings of the National Academy of Sciences*, *93*(20), 10858–10863. https://doi.org/10.1073/pnas.93.20.10858

Stopczynski, R. E., Poloskey, S. L., & Haber, S. N. (2008). Cell proliferation in the striatum during postnatal development: Preferential distribution in subregions of the ventral striatum.

*Brain Structure & Function*, *213*(0), 119–127. https://doi.org/10.1007/s00429-008-0185-1

Struhl, K. (1998). Histone acetylation and transcriptional regulatory mechanisms. *Genes & Development*, *12*(5), 599–606.

Stühmer, T., Anderson, S. A., Ekker, M., & Rubenstein, J. L. R. (2002). Ectopic expression of the Dlx genes induces glutamic acid decarboxylase and Dlx expression. *Development (Cambridge, England)*, *129*(1), 245–252.

Stühmer, T., Puelles, L., Ekker, M., & Rubenstein, J. L. R. (2002). Expression from a Dlx Gene Enhancer Marks Adult Mouse Cortical GABAergic Neurons. *Cerebral Cortex*, *12*(1), 75–85. https://doi.org/10.1093/cercor/12.1.75

Tan, X., & Shi, S.-H. (2013). Neocortical neurogenesis and neuronal migration. *Wiley Interdisciplinary Reviews. Developmental Biology*, *2*(4), 443–459. https://doi.org/10.1002/wdev.88

The ENCODE Project Consortium. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, *447*(7146), 799–816. https://doi.org/10.1038/nature05874

The ENCODE Project Consortium. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, *489*(7414), 57. https://doi.org/10.1038/nature11247

Thomas, B. L., Tucker, A. S., Qui, M., Ferguson, C. A., Hardcastle, Z., Rubenstein, J. L., & Sharpe, P. T. (1997). Role of Dlx-1 and Dlx-2 genes in patterning of the murine dentition. *Development (Cambridge, England)*, *124*(23), 4811–4818.

Tie, F., Banerjee, R., Conrad, P. A., Scacheri, P. C., & Harte, P. J. (2012). Histone Demethylase UTX and Chromatin Remodeler BRM Bind Directly to CBP and Modulate Acetylation of Histone H3 Lysine 27. *Molecular and Cellular Biology*, *32*(12), 2323–2334. https://doi.org/10.1128/MCB.06392-11

Turrero García, M., & Harwell, C. C. (2017). Radial glia in the ventral telencephalon. *FEBS Letters*, *591*(24), 3942–3959. https://doi.org/10.1002/1873-3468.12829

Ucar, D., Hu, Q., & Tan, K. (2011). Combinatorial chromatin modification patterns in the human genome revealed by subspace clustering. *Nucleic Acids Research*, *39*(10), 4063–4075. https://doi.org/10.1093/nar/gkr016

Vastenhouw, N. L., & Schier, A. F. (2012). Bivalent Histone Modifications In Early Embryogenesis. *Current Opinion in Cell Biology*, *24*(3), 374–386. https://doi.org/10.1016/j.ceb.2012.03.009

Verdin, E., & Ott, M. (2015). 50 years of protein acetylation: From gene regulation to epigenetics, metabolism and beyond. *Nature Reviews Molecular Cell Biology*, *16*(4), 258–264. https://doi.org/10.1038/nrm3931

Voigt, P., Tee, W.-W., & Reinberg, D. (2013). A double take on bivalent promoters. *Genes & Development*, *27*(12), 1318–1338. https://doi.org/10.1101/gad.219626.113

von Bartheld, C. S., Bahney, J., & Herculano-Houzel, S. (2016). The search for true numbers of neurons and glial cells in the human brain: A review of 150 years of cell counting. *The Journal of Comparative Neurology*, *524*(18), 3865–3895. https://doi.org/10.1002/cne.24040

von Schimmelmann, M., Feinberg, P. A., Sullivan, J. M., Ku, S. M., Badimon, A., Duff, M. K., … Schaefer, A. (2016). Polycomb repressive complex 2 (PRC2) silences genes responsible for neurodegeneration. *Nature Neuroscience*, *19*(10), 1321–1330. https://doi.org/10.1038/nn.4360

Wang, Y., Dye, C. A., Sohal, V., Long, J. E., Estrada, R. C., Roztocil, T., … Rubenstein, J. L. R. (2010). Dlx5 and Dlx6 Regulate the Development of Parvalbumin-Expressing Cortical Interneurons. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(15), 5334–5345. https://doi.org/10.1523/JNEUROSCI.5963-09.2010

Wang, Z., Zang, C., Rosenfeld, J. A., Schones, D. E., Barski, A., Cuddapah, S., … Zhao, K. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. *Nature Genetics*, *40*(7), 897–903. https://doi.org/10.1038/ng.154

Warren, N., Caric, D., Pratt, T., Clausen, J. A., Asavaritikrai, P., Mason, J. O., … Price, D. J. (1999). The transcription factor, Pax6, is required for cell proliferation and differentiation in the developing cerebral cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, *9*(6), 627–635. https://doi.org/10.1093/cercor/9.6.627

Wiles, E. T., & Selker, E. U. (2017). H3K27 methylation: A promiscuous repressive chromatin mark. *Current Opinion in Genetics & Development*, *43*, 31–37. https://doi.org/10.1016/j.gde.2016.11.001

Woo, Y. H., & Li, W.-H. (2012). Evolutionary conservation of histone modifications in mammals. *Molecular Biology and Evolution*, *29*(7), 1757–1767. https://doi.org/10.1093/molbev/mss022

Wyrick, J. J., & Parra, M. A. (2009). The role of histone H2A and H2B post-translational modifications in transcription: A genomic perspective. *Biochimica Et Biophysica Acta*, *1789*(1), 37–44. https://doi.org/10.1016/j.bbagrm.2008.07.001

Xiao, B., Wilson, J. R., & Gamblin, S. J. (2003). SET domains and histone methylation. *Current Opinion in Structural Biology*, *13*(6), 699–705.

Xu, C., Funahashi, Y., Watanabe, T., Takano, T., Nakamuta, S., Namba, T., & Kaibuchi, K. (2015). Radial Glial Cell-Neuron Interaction Directs Axon Formation at the Opposite Side of the Neuron from the Contact Site. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(43), 14517–14532. https://doi.org/10.1523/JNEUROSCI.1266-15.2015

Yu, M. (2011). Functional analysis of *cis*-regulatory elements I56i, I56ii and I12b that control *Dlx* expression in the developing forebrain of mouse and zebrafish (PhD dissertation, University of Ottawa (Canada)).

Zentner, G. E., & Scacheri, P. C. (2012). The Chromatin Fingerprint of Gene Enhancer Elements. *The Journal of Biological Chemistry*, *287*(37), 30888–30896. https://doi.org/10.1074/jbc.R111.296491

Zerucha, T., Muller, J.-P., Chartrand, N., & Ekker, M. (1997). Cross-interactions between two members of the Dlx family of homeobox-containing genes during zebrafish development. *Biochemistry and Cell Biology*, *75*(5), 613–622. https://doi.org/10.1139/o97-076

Zerucha, T., Stühmer, T., Hatch, G., Park, B. K., Long, Q., Yu, G., … Ekker, M. (2000). A highly conserved enhancer in the Dlx5/Dlx6 intergenic region is the site of cross-regulatory

interactions between Dlx genes in the embryonic forebrain. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *20*(2), 709–721.

Zheng, C., & Hayes, J. J. (2003). Structures and interactions of the core histone tail domains. *Biopolymers*, *68*(4), 539–546. https://doi.org/10.1002/bip.10303

Zhou, P. (2017). Phenotype characterization of mice with targeted deletions of *Dlx* enhancers (M.Sc. dissertation, University of Ottawa (Canada)).