Activity of *Dlx* transcription factors in regulatory cascades underlying vertebrate forebrain development

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies in partial fulfillment of the requirements for the Master of Science degree
Ottawa-Carleton Institute of Biology
University of Ottawa

Thèse soumise à la Faculté des Études Supérieures et Postdoctorales
En vue de l'obtention de la Maitrise ès Sciences
L'institut de Biologie d'Ottawa-Carleton
Université d'Ottawa

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

ANB – Anterior border of the neural plate

A/P – Anterior-posterior

Arx – Arista-less homeobox

Ascl – achaete-scute complex like

BAC – Bacterial artificial chromosome

bHLH – Basic helix-loop-helix

BMP – Bone morphogenetic protein

ChIP – Chromatin immunoprecipitation

CNE – Conserved non-coding element

CNS – Central nervous system

CRE – *Cis*-regulatory element

Dlx – Distal-less homeobox

Dpf – Days post fertilization

D/V – Dorsal-ventral

eGFP – Enhanced green fluorescence protein

Evo-Devo – Evolutionary developmental biology

Fez – Forebrain embryonic zinc-finger

FGF – Fibroblast growth factor

GABA – Gamma-Aminobutyric acid

Gad – Glutamic acid decarboxylase

GRN – Gene regulatory network

HD - Homeodomain

Hh - Hedgehog

Hox - Homeobox

Hpf – Hours post fertilization

Irx – Iroquois homeobox

Lhx – LIM-homeodomain transcription factors

LGE – Lateral Ganglionic eminence

Mash – Mouse achaete-scute homolog-

MDO – Mid –diencephalon organizer

MGE – Medial ganglionic eminence

MO – Morpholino oligonucleotide

Ngn - Neurogenin

NPC – Neural progenitor cell

NSC – Neural stem cell

OPC – Oligodendrocyte precursor cell

RA – Retinoic acid

REST – RE-1 silencing factor

sFRP – Secreted frizzled-related protein

Shh – Sonic hedgehog

SVZ – Sub-ventricular

URE – Upstream regulatory element

VZ – Ventricular zone

ZLI – Zona limitans intrathalamica

ABSTRACT

The temporal and spatial patterning that underlies morphogenetic events is controlled by gene regulatory networks (GRNs). These operate through a combinatorial code of DNA – binding transcription factor proteins, and non – coding DNA sequences (*cis*-regulatory elements, or CREs), that specifically bind transcription factors and regulate nearby genes. By comparatively studying the development of different species, we can illuminate lineage – specific changes in gene regulation that account for morphological evolution.

The central nervous system of vertebrates is composed of diverse neural cells that undergo highly coordinated programs of specialization, migration and differentiation during development. Approximately 20% of neurons in the cerebral cortex are GABAergic inhibitory interneurons, which release the neurotransmitter gamma-aminobutyric acid (GABA). Diseases such as autism, schizophrenia and epilepsy are associated with defects in GABAergic interneuron function. Several members of the *distal-less homeobox* (*Dlx*) transcription factor family are implicated in a GRN underlying early GABAergic interneuron development in the forebrain.

I examined the role played by orthologous *dlx* genes in the development of GABAergic interneurons in the zebrafish forebrain. I found that when *ascl1a* transcription factor is down-regulated through the micro-injection of translation – blocking morpholino oligonucleotides, *Dlx* gene transcription is decreased in the diencephalon, but not the telencephalon. Similarly, *gad1a* transcription is also decreased in this region for these morphants. As *gad1a* encodes an enzyme necessary for the production of GABA, these genes are implicated in a cascade underlying GABAergic interneuron development in the diencephalon.

RÉSUMÉ

Au cours du développement embryonnaire, les événements spatio-temporels impliqués dans la morphogenèse sont contrôlés par des réseaux de régulations géniques (RRG). Cela s'opère par l'intermédiaire d'un code combinatoire impliquant des protéines de types facteur de transcription qui se lient à l'ADN et des séquences d'ADN non-codante (éléments cis-régulateurs ou ECR) qui fixent spécifiquement ces facteurs de transcription pour réguler les gènes à proximité. Par l'intermédiaire d'études comparatives du développement de différentes espèces, il est possible de mettre en lumière comment des changements de régulations de gènes auraient pu être impliqués dans la genèse de nouvelles formes au cours de l'évolution.

Le système nerveux central des vertébrés est composé de cellules neurales diverses qui suivent un programme hautement coordonné de spécialisation, de migration et de différenciation pendant le développement. Au sein du cortex cérébral, environ 20% des neurones sont des interneurones GABAergiques inhibiteurs qui sécrètent comme neurotransmetteur l'acide gamma-aminobutyrique (GABA). De plus, plusieurs membres de la famille des facteurs de transcription *distal-less homeobox* (*Dlx*) sont connus pour être impliqués dans un RRG ayant un rôle central dans le développement précoce des interneurones GABAergiques du cerveau antérieur.

J'ai ici examiné l'implication des gènes orthologues *dlx* dans le développement des neurones GABAergiques du poisson zèbre. J'ai démontré que la réduction éxperimental de l'expression *ascl1a* cause un diminuation de transcription des gènes *Dlx* dans la partie diencéphalique et non télencéphalique du cerveau antérieur. Aussi, dans les mêmes morphants, la transcription de *gad1a* est aussi diminué dans cet région. Car gad1a code une enzyme nécessaire pour la production de GABA, ces gènes sont impliqués dans une cascade étant à la base du développement d'interneuron de GABAergic dans le diencephalon.

STATEMENT OF CONTRIBUTIONS

Jacob Pollack did all work and analysis presented as such in this thesis.

ACKNOWLEDGEMENTS

I would like to thank Marc Ekker for providing me with this opportunity to do research in his lab and for always having time to provide invaluable feedback and council. Mélanie Debiais-Thibaud patiently mentored me and taught me the ways of molecular and developmental biology. Her love of Evo-Devo is contagious, and I greatly appreciate the time and effort she put into making this degree a fruitful learning experience for me. Eglantine Heude helped me properly design experiments, optimize protocols, use imaging software, analyze results, write coherently, and to keep smiling during times of potential insanity. I would like to thank Gary Hatch and Vishal Saxena who were always happy to share their wealth of knowledge and help me in the lab and fishroom, respectively. I appreciate all the Ekker lab members for sharing ideas and making the lab environment friendly, helpful and conducive to learning. Finally, I would like to thank my family for the support and encouragement they've shown me.

INTRODUCTION

1.1.1 Gene regulatory networks (GRNs):

A central goal in biology is to characterize the relationship between genotype and phenotype of life forms. How has a simple and highly conserved genetic code of inheritance given rise to such an overwhelming diversity of organisms? Prior to the recent capacity of drawing mass genomic sequence comparisons between species, morphological variation was thought to derive from lineage-specific mutations in protein-coding regions; the discrete phenotypic effects of which were subject to natural selection. However, the advances in genomic sequencing have revealed a generally high level of gene conservation between even distantly related species. This observation rendered a strict causal relationship between morphological and molecular divergence more tenuous. In addition, the discovery of abundant conserved non-protein-coding regulatory sequences indicated a more central role of gene regulation in the creation of morphological complexity. Rather than being autonomous and isolated, the expression of a given gene is contingent on larger gene regulatory networks (GRNs).

Generally, a GRN is made up of genes coding for DNA-recognizing proteins (transcription factors), non-coding DNA sequences which are bound by proteins and regulate gene expression (*cis*-regulatory elements, CREs), and genes whose products are not transcription factors. GRNs have two broad components: (1) Genes and their accompanying regulatory sequences are referred to as *nodes*, and comprise the genomic component, and (2) transcription factors producing regulatory inputs to these nodes make up the regulatory state component (Davidson, 2006). Once components of a network have become characterized through experimentation, they can be mapped out and the network 'regulatory architecture' determined.

Far from being static, GRNs can arise transiently and specific to cell type/location, stage of cell cycle, and developmental time point.

1.1.2 GRNs in developmental processes

During development, the thousands of genes in a species' genome must be correctly expressed in space and time, in an increasingly complex body. Cells must "sense" their context in order to express the necessary subset of genes. A developmental GRN (1) receives inputs from an initial regulatory state, (2) undergoes a specific cascade of node activation, and (3) generates outputs in the form of a final regulatory state. Development may thus be seen as the continuous generation of successive and regionally multiplying regulatory states, which underlie spatial gene expression, specification and differentiation of cells, and morphogenesis.

Spatial gene expression is a phenomenon occurring from subcellular to organismal biological organization. In bilaterians, early gene expression patterns define anterior/posterior and dorsal/ventral axes, and metameric segmentation. The formation of later expression patterns dictates the spatial organization of body parts, and even later ones define smaller and more specific corporal features. By expressing new sets of transcription factors in a defined group of cells, subdivisions are made within a larger group of cells and regional specification occurs. The mechanisms at work during this process invariably include *CREs,* which regulate genes encoding transcription factors involved in regional specification; locally -defined regulatory states provide inputs to cell populations to designate specific developmental functions (Davidson, 2006). Fully differentiated cell populations have reached a terminal developmental regulatory state, and express subsets of genes that encode the characteristic attributes of each cell type. The entire

assembly of GRNs employed in an organism's life is encoded in its DNA, and makes up its 'regulatory genome'.

1.1.3 GRNs and evolutionary developmental biology

The general approach of evolutionary developmental biology, or Evo-Devo, is to investigate how the evolution of novel morphology is related to inherited changes in developmental processes. Gene regulatory networks play an important role in all levels of biological organization, from early body plan organization to terminal cell differentiation. Inter-species morphological variation can be attributed to inherited differences in developmental GRNs, encoded in each species' regulatory genome. In other words, evolution of organismal form arises primarily through lineage-specific changes in developmental GRN architecture. Mutations in a CRE sequence can alter its affinity to bind transcription factors, and thus induce differential gene regulation. Larger mutations can completely remove an enhancer from its position near a gene. This shift in input/output states of a single node can bring about new dynamics within a larger GRN. Furthermore, should this result in different spatial patterning of the embryo, or any other significant difference in development, a novel phenotype could be induced. Finally, the new phenotype would be subject to natural selection, and if beneficial (or at least not too deleterious), the CRE mutation may be fixed in the population.

One can imagine a hypothetical example: A mutation occurring in the CRE of a gene responsible for inducing the proliferation of early limb bud cells causes the gene to be expressed for a longer time during development. This results in a greater proliferation of embryonic limb bud cells and the subsequent formation of longer limbs. As this bestows a fitness advantage upon the individual, the trait and its underlying mutation are thereafter maintained in a population.

Lineage-specific mutations in gene regulation may result in heterochrony, or different *timing* of developmental gene expression between species. Similarly, heterotopy occurs when there are lineage-specific changes in *where* a gene is expressed during development. Both these instances can result in altered developmental processes and therefore yield morphological divergence.

In addition to regulatory sequence mutation as a mechanism for developmental changes in evolution, the occurrence of gene and genome duplication events of developmental transcription factor genes produces new 'ready-made' nodes that may be co-opted into new or old GRNs. A comparative analysis of developmental GRNs between species can yield important insights on the evolution of organismal form. This can be done experimentally through functional analysis of developmental GRNs. As the interaction of transcription factors and CREs plays a central role in the regulatory genomic code, it is necessary to describe their function in greater detail.

We can approach evolutionary questions by drawing inter-species comparisons of important developmental GRNs; changes within which might underlie differences in adult morphology. Furthermore, inter-species comparative analysis of conserved enhancer sequences can reveal mutated sites leading to altered regulatory activity and developmental function, inducing evolution.

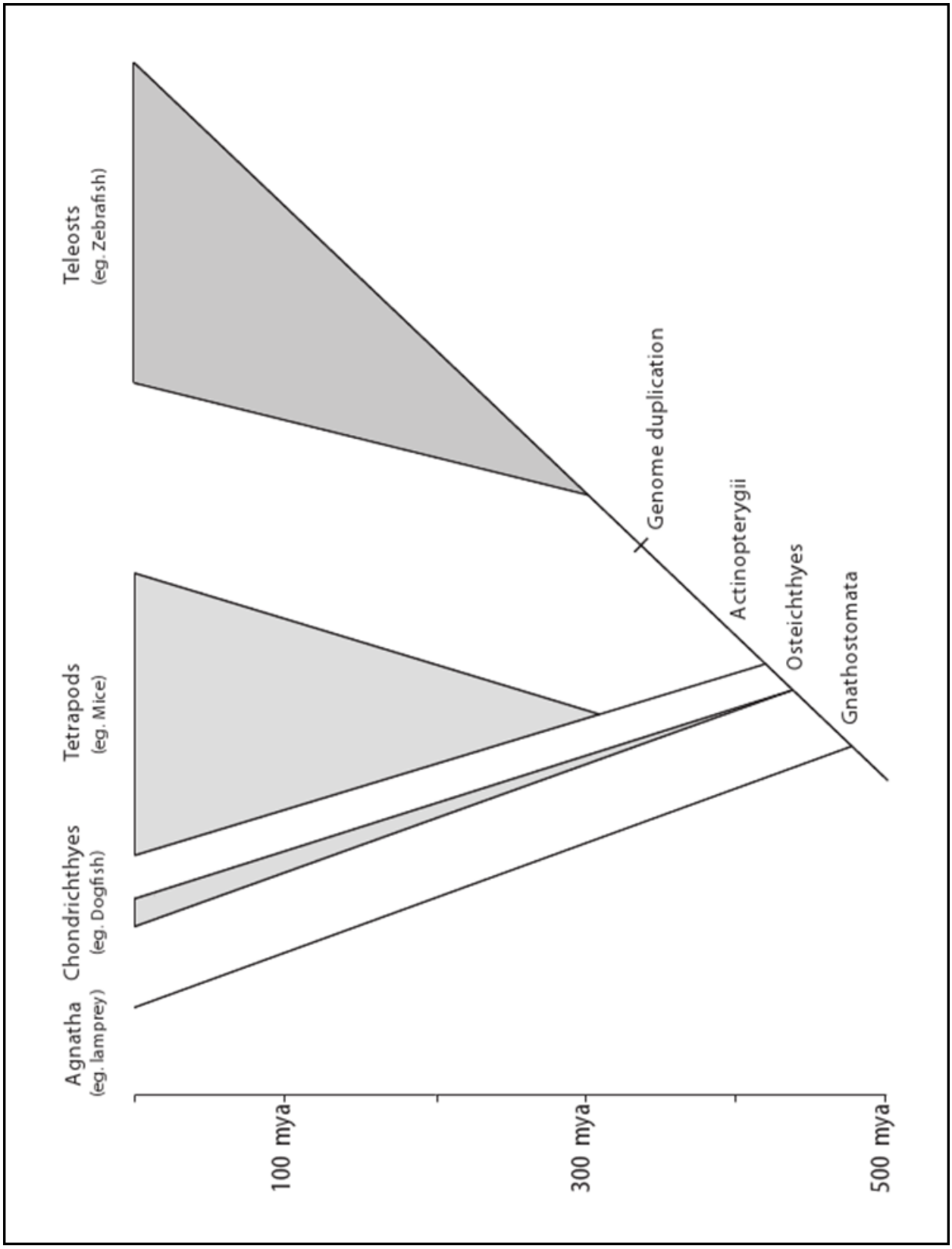
1.1.4 Comparative genomics and the evolution of CREs

In comparative genomics, homologous DNA sequences are compared between different species to identify regions that are functionally important. Due to selective pressure, it is assumed that these regions will experience a lower mutation rate than those lacking a functional or protein-coding role (Kimura, 1983). Conserved non-coding elements (CNEs) are untranscribed sequences of DNA that have relatively low mutation rates, and it is thought that

many CNEs contain CREs (Pennachio *et al*., 2006). Although CREs are not protein-coding DNA sequences, the specific binding of transcription factors endows them a functional role in the genome, and so they can experience selective pressure. Therefore, comparative genomics can be used to identify putative enhancer sequences conserved between species. Species-specific variations in the expression of developmental genes are the basis for large intra-species morphological differences. As these differences in expression reflect changes in gene regulation, it follows logically that the study of CREs near developmental genes may provide information on how molecular changes affect the evolution of organismal form. Interestingly, there is not always a strict relationship between the CRE sequence conservation and regulatory function. Although many highly conserved homologous CREs activate gene expression in the same tissues or developmental time points in their respective species, there are cases in which their regulatory functions diverge (Dickmeis and Müller, 2005).

The separation between bony (Osteichthyes), and cartilaginous (chondrichthyes), vertebrate lineages occured an estimated 450 million years ago (mya) (Venkatesh *et al*, 2007). Approximately 20 million years later Osteichthyes diverged into Actinopterygii (including the teleost fish), and Sarcopterygii (including the tetrapods) (**Figure 1**). There is considerable conservation of developmental gene and CRE sequences, as well as synteny (the physical co-localization of genes on a chromosome), between these three broad vertebrate groups. Since mammalian and teleost homologous CREs are more recently diverged, will they have more similar regulatory functions than those of cartilaginous fish? Conversely, has the extra genome duplication event in teleosts resulted in a 'relaxation' of selective pressure on CRE sequences, promoting the emergence of new regulatory functions in this lineage? These types of questions can be approached through the comparative analysis of CRE function. The degree to

Figure 1: Phylogenetic tree of extant vertebrate groups. This tree shows how four extant vertebrate groups diverged from common ancestors over the last several hundred million years. Above each branch is the group name and an example of one species in the group: agnathans (lamprey), chondrichthyes (dogfish), tetrapods (mouse), and teleost (zebrafish). Name of common ancestor is written below each branching point. Grey regions represent diversification within a group. The lineage leading to teleosts is thought to have had an extra genome duplication event. Branch lengths are done to scale based on fossil records (Venkatesh *et al*., 2007).


Teleosts
(eg. Zebrafish)
Tetrapods
(eg. Mice)
Chondrichthyes
(eg. Dogfish)
Agnatha
(eg. lamprey)
Genome duplication
Actinopterygii
Osteichthyes
Gnathostomata
100 mya
300 mya
500 mya

which CRE regulatory function has been conserved between vertebrate lineages can be investigated experimentally.

One approach to CRE functional analysis is to introduce into the embryonic genome a sequence containing the CRE of interest upstream of a reporter gene (like *eGFP* or *LacZ*) driven by a minimal promoter. The larvae are thereafter examined for regulation of the given reporter gene. Two central assumptions of CRE functional analysis are: (1) Endogenous transcription factors will interact with this introduced regulatory region as if it were an endogenous target, and (2) upon being bound by endogenous transcription factors, an introduced CRE will activate a nearby reporter gene similarly to the gene(s) it regulates endogenously. A different approach to CRE functional analysis is to induce mutations in, or delete, specific putative CRE sites in an organism and examine the effect this has on expression of the regulated gene. For both these techniques transgenic lines can be made, in which every cell in an individual organism contains the introduced or altered CRE sequence. Although transgenesis allows for a more precise study of CRE function at any developmental stage, transient expression of an introduced CRE-reporter plasmid may be adequate to draw some general conclusions at early embryonic time points (before the plasmid becomes too diluted during growth).

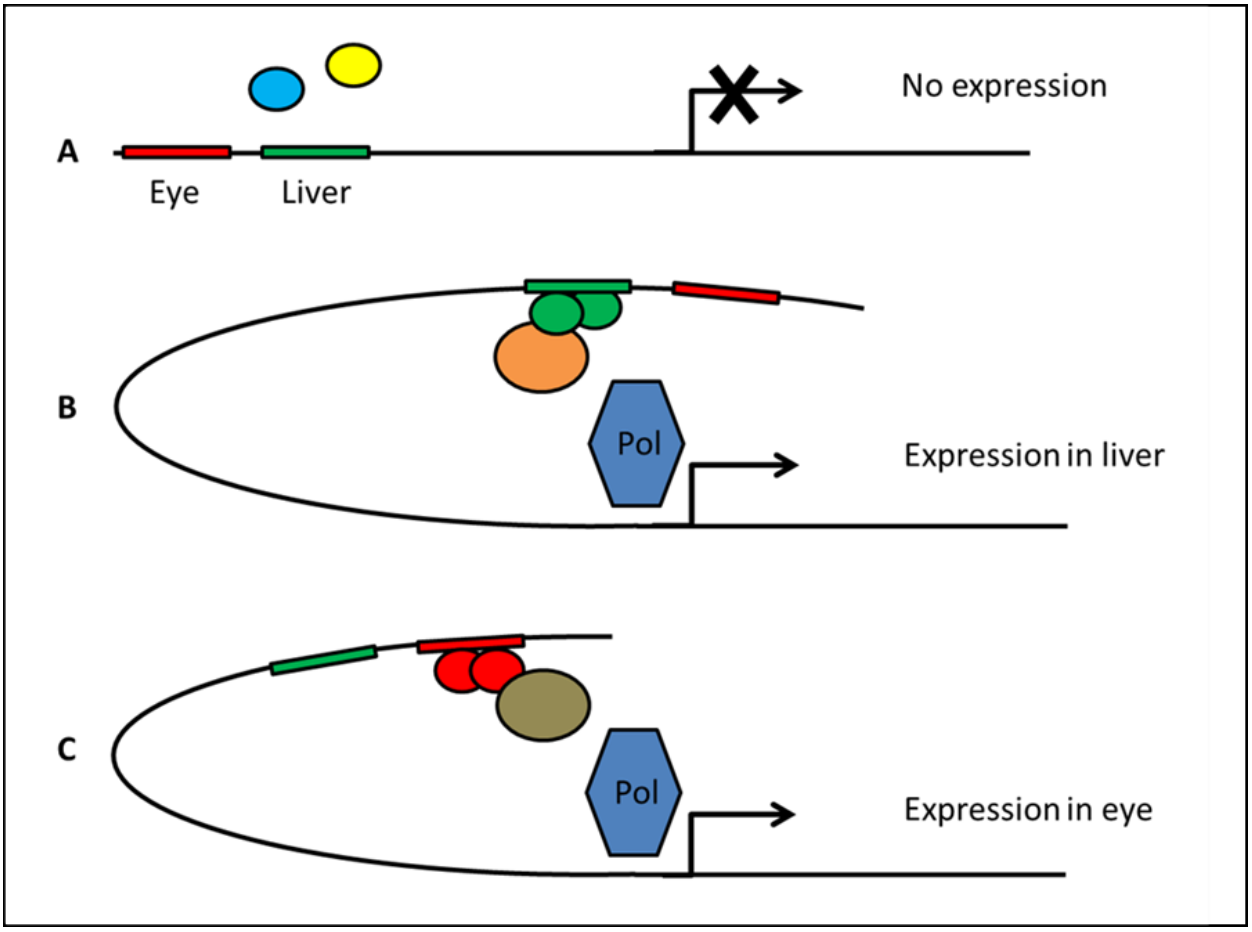
The vertebrarte central nervous system (CNS) is composed of a wide variety of specialized neural cell subtypes. Experimental study of early CNS development can yield information on how the specification, differentiation and migration of the diversity of cell types is regulated at the molecular level. The present work focuses on illuminating these mechanisms in the brain, which, due to its complex organisation at both cellular and tissue levels, requires further study.

1.1.5 The CRE- mediated combinatorial code

Although several classes of functional non-coding regulatory elements exist, *cis*-regulatory elements are central in developmental information processing. They conditionally control spatial and temporal expression of genes located on the same chromosome. Within their sequences are contained densely packed clusters of short sequence motifs, specifically recognized and bound by a diversity of transcription factors. The regulatory output of a CRE is dependent on the combination and proximity of transcription factors bound at a given time. "Enhancers" are CREs that positively regulate genes, and, being positioned up to several kilobases away from the basal transcription apparatus, must form loops to gain proximity to where transcription machinery is recruited (**Figure 2)**. Because of this looping, the orientation of the enhancer sequence is irrelevant. CREs may be located 5' or 3' of a gene, or in introns. Enhancers operate in two ways: (1) 'Billboard' enhancers consist of several transcription binding sites which act independently from one another to affect the transcription of a target gene (Kulkarni and Arnosti, 2003), and (2) the enhanceosome model predicts that basal transcription machinery and transcription factors are specifically arranged and act in cohort to regulate transcription (Panne *et al.*, 2007). One enhancer element may activate the expression of several genes, and conversely, several enhancers may activate one gene.

"Repressor" CREs may prohibit the expression of a gene, while "insulators" act to restrict nearby enhancer activity to specific genes. The RE-1 silencing factor (REST) is an example of a transcriptional repressor, which is activated in non-neural cells to prevent the transcription of neural specific genes. REST binds to repressor elements and induces histone acetylation, which alters the chromatin structure and inhibits transcription of genes. Many regulatory genes each have temporally and spatially distinct functions throughout the life of an organism. These genes

Figure 2: **Cartoon representing tissue – specific regulatory function of enhancer *cis*-regulatory elements (CREs).** (A) Upstream of its transcription start site, an hypothetical gene has two enhancer elements that conditionally activate transcription in eye (red bar) and liver (green bar) tissues. In the absence of transcription factor that target each enhancer (blue and yellow circles represent non-binding transcription factors), a protein-DNA complex is unable to form and therefore transcription machinery is not recruited to begin gene expression. (B) In the presence of transcription factors that are expressed in the liver (green circles) and target the liver –specific enhancer, a protein-DNA complex is formed that also binds cofactors (orange). This results in the conformational change in the DNA molecule, allowing transcriptional machinery to be recruited to the transcription start site, and the activation of gene expression in the liver. (C) Similarly to in (B), when eye –specific transcription factors are present, a protein-DNA complex is formed with co-factors (brown) that causes a conformational change in the DNA molecule and recruits transcription machinery, activating gene expression in the eye. Arrows- Gene transcription start site, Pol- polymerase transcription enzyme.


A
No expression
Eye
Liver
B
Pol
Expression in liver
C
Pol
Expression in eye

must therefore respond to various regulatory states. The solution is for a gene to have several enhancers, each activated by a combination of transcription factors unique to a specific regulatory state. The specific regulatory action of each enhancer of a given gene may be determined experimentally. When inserted into an embryo, an isolated enhancer sequence is able to drive a reporter gene to recapitulate some, if not all, of the endogenous gene expression pattern. During development, cascades of gene regulation occur in which initial 'upstream' transcription factors bind enhancers, activating the expression of other transcription factors, which in turn bind to enhancers of 'downstream' genes. In this way, a single transcription factor's signal may be amplified and lead to the expression of diverse genes involved in different developmental processes (Shirasaki and Pfaff, 2002).

1.2.1 Overview of vertebrate brain development and structure

The ectoderm is the most peripheral of the three early embryonic germ layers in vertebrates. The future nervous system develops from the neuroectoderm- the dorsal region of this germ layer. It separates early from the more ventrolateral general ectoderm, which instead becomes the future epidermal skin and its associated structures (Mueller and Wulliman, 2003). Through the process of neurulation, the general ectoderm engulfs the central part of the neuroectoderm, or neural plate, separating the two and bringing the latter deeper into the embryo. The neural plate develops into a hollow, cerebrospinal fluid-filled neural tube. The dorsal opening created by the engulfment of the neural plate is enclosed by the general ectoderm. The anterior neural tube develops into the brain, and through vertical constrictions is subsequently divided into a series of vesicles along the anterior-posterior axis. The first division separates a combined forebrain (prosencephalon) and midbrain (mesencephalon) vesicle, from the hindbrain

(rhombencephalon) vesicle. Next, there is a three-vesicle stage with prosencephalon, mesencephalon and rhombencephalon vesicles. The prosencephalon is then divided into the telencephalon and diencephalon, and the rhombencephalon is divided into the metencephalon and myelencephalon, totaling five brain subdivisions (**Figure 3)**. Along the dorsal-ventral axis, the neural tube is divided into four longitudinal zones of the roof (dorsal), alar, basal and floor (ventral) plates- zones thought to extend anteriorly into the forebrain (Hauptmann and Gerster, 2000). The neuromeric model proposes that the vertebrate brain is further made up of smaller transverse developmental 'units', called neuromeres, between which cellular migration is limited (Pueller and Rubenstein, 2003). Thus, the embryonic brain is divided into a 'checkerboard' of transverse and longitudinal zones. It must be noted that the neural tube undergoes two flexures during development; a cephalic one between the prosencephalon and mesencephalon, and a cervical one between the rhombencephalon and spinal cord. Each flexure causes the neural tube to form an upside-down 'U' shape. This bending invites the mis-identification of the correct anterior-posterior and dorsal-ventral axes (**Figure 4)**.

The vertebrate prosencephalon, (the telencephalon and diencephalon), is of primary concern in the present work. Neuromeres within this region are called prosomeres and there is evidence for at least four of these located in the diencephalon. These are (anteriorly to posteriorly) the hypothalamus, prethalamus, the thalamus and pretectum (Mueller and Wullimann, 2002b). The zona limitans intrathalamica (ZLI) is an embryonic signaling center located between the prethalamus and thalamus (Scholpp *et al.,* 2006). The diencephalon develops into the thalamus, subthalamus, hypothalamus and epithalamus. In general, these structures sit upon the brain stem and perform sensory processes, motor function, autonomic activities and sleep regulation (Roth and Wulliman, 2001). The telencephalon develops into the

Figure 3: Cartoon depicting archetypal embryonic vertebrate brain patterning. After neurulation, the anterior neural tube pinches into the distinct segments of the prosencephalon (forebrain), mesencephalon (midbrain), and rhombencephalon (hindbrain). The prosencephalon then divides into the telencephalon and diencephalon, and the rhombencephalon divides into the metencephalon and myelencephalon.

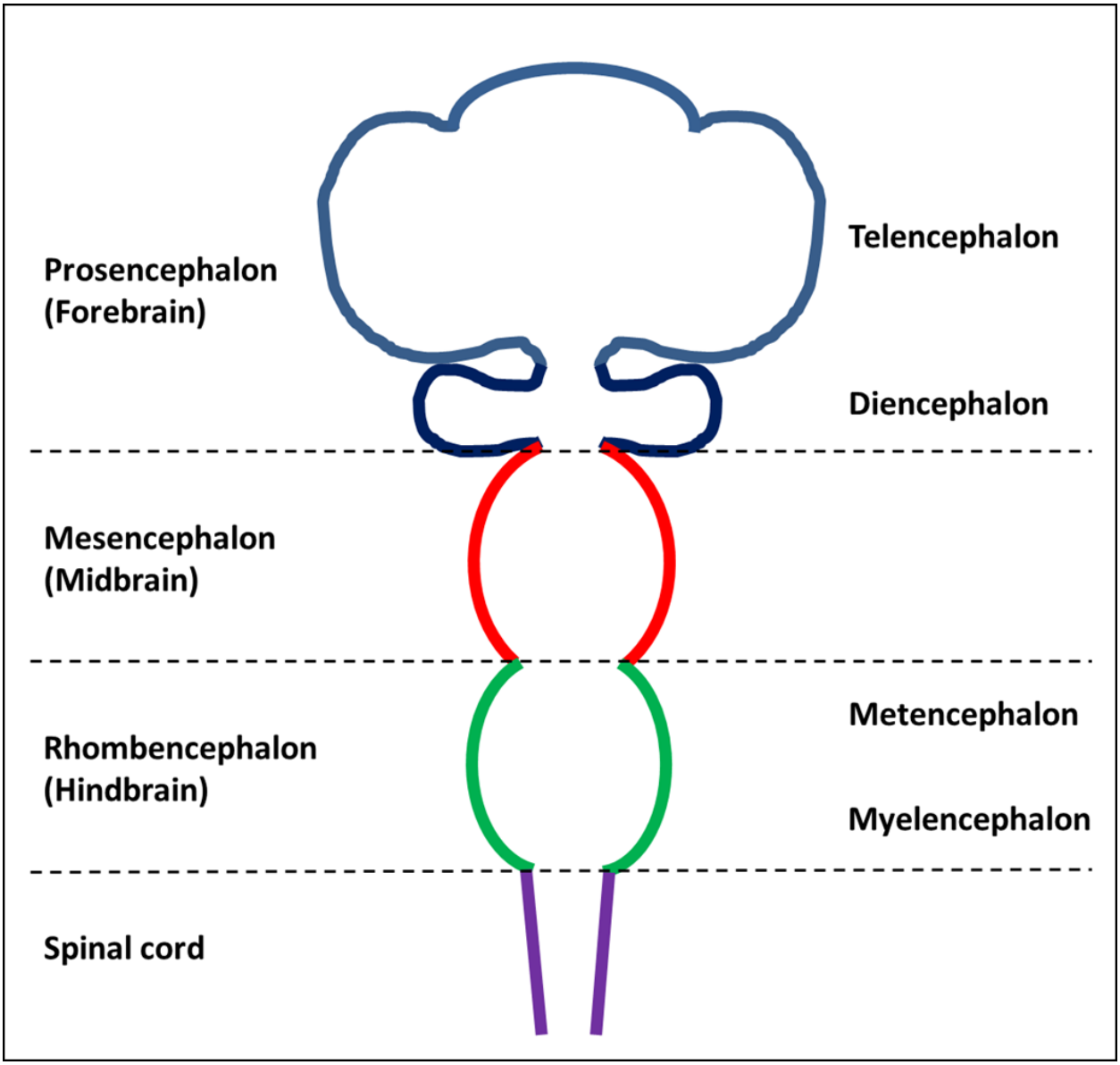
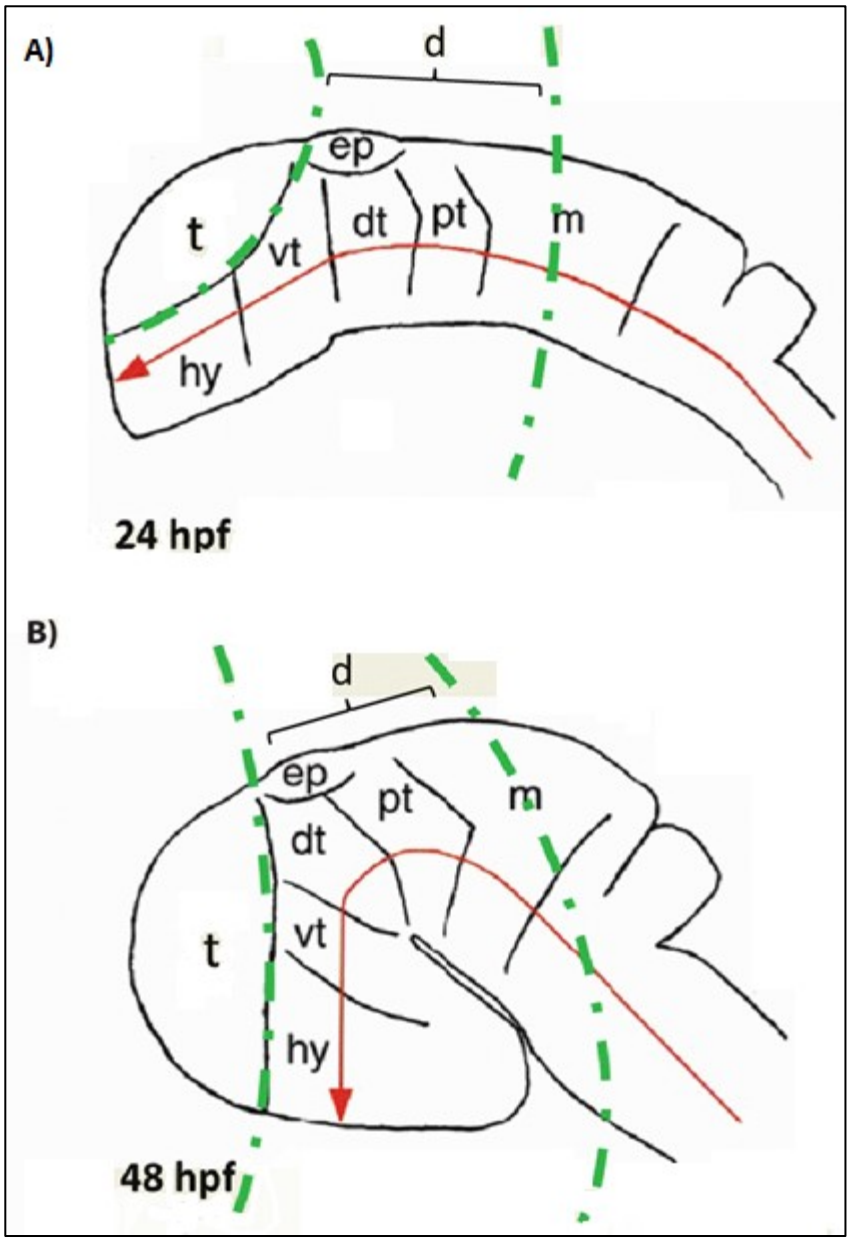

Prosencephalon
(Forebrain)
Telencephalon
Diencephalon
Mesencephalon
(Midbrain)
Rhombencephalon
(Hindbrain)
Metencephalon
Myelencephalon
Spinal cord

Figure 4: Diagram showing lateral sagittal view of early brain region formation and anterior neural neural flexure in zebrafish. Positioning of forebrain regions changes between (A) 24 hpf and (B) 48 hpf with respect to the dorsoventral and anteroposterior axes. Red lines illustrate the degree to which the anterior neural flexure alters the orientation of the forebrain with respect to the midbrain. Left green dotted lines separate the telencephalon from the diencephalon in the forebrain, and the right green dotted lines separate the forebrain from the midbrain. t -telencephalon, d -diencephalon, hy –hypothalamus, vt –ventral thalamus, dt –dorsal thalamus, ep –epiphysis, pt –pretectum, m –midbrain.

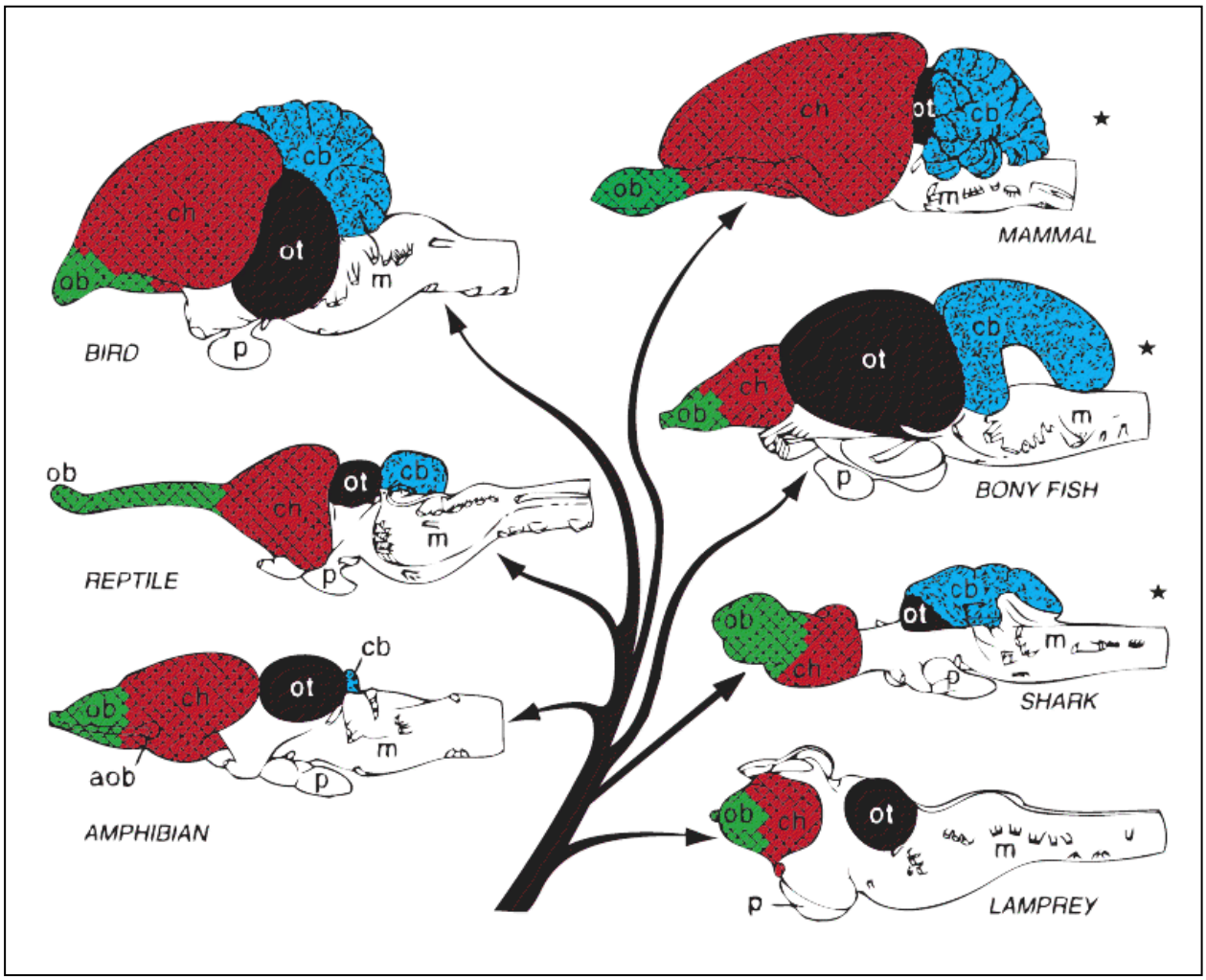

A)
d
ep
t
vt
dt
pt
m
hy
24 hpf
B)
d
ep
pt
m
dt
t
vt
hy
48 hpf

dual-lobed cerebrum. The dorsal telencephalon, or pallium, is precursor to the cerebral cortex, and the ventral telencephalon, or subpallium, gives rise to the basal ganglia (Roth and Wulliman, 2001). The functions of the cerebrum include voluntary movement, sensory processing, communication, learning, memory and olfaction (Roth and Wulliman, 2001).

1.2.2 Comparative approaches to brain evolution

A comparative look at the vertebrate brain can reveal both deep homologies, as well as extreme divergences between distantly related species. For example, all jawed vertebrates have the same number of brain divisions. Furthermore, the targeted brain regions of the ascending spinal neural projections are the same among lamprey, bony and cartilaginous fish, amphibians, reptiles, birds and mammals (Ebbeson, 1980). The relative division size, overall brain size and number of neural centers, however, are highly variable between vertebrates (**Figure 5)**. A particularly extreme case of morphological divergence is the optic tectum of teleost fish, which is enormous compared to that of both cartilaginous fish and mammals (Northcutt, 2002). Furthermore, evagination (in mammals) and eversion (in teleosts) are two contrasting patterns of neural tube folding, yielding major structural differences between mammals and teleosts. Most vertebrate brains undergo evagination, where the telencephalon forms as two telencephalic hemispheres surrounding a central ventricle (**Figure 6)**. Eversion, however, is characterized by a lateral rolling-out of the pallium, ending with the ventricular surface on the outside of the pallium (Niewenhuys and Meek, 1990). Despite this fundamental difference in neural tube folding, we can still draw comparisons between highly analogous brain structures. Of course, these large-scale differences in brain morphology arise from differences in morphogenetic events

Figure 5: Phylogenic tree of extant vertebrate lineages showing lateral drawings of brains. Despite considerable variation in the relative sizes of individual brain regions between lineages, the basic brain plan is highly conserved. Selected brain regions that are presumably homologous are colour coded. Green – (ob) olfactory bulb, red – (ch) cerebral hemispheres, black – (ot) optic tectum, blue – (cb) cerebellum, p – pituitary, m – medulla oblongata. Stars indicate lineages between which experimental comparisons are made in this work. Branch length and brain sizes are not to scale. Anterior is to the left, and dorsal is up (adapted from Northcutt, 2002).

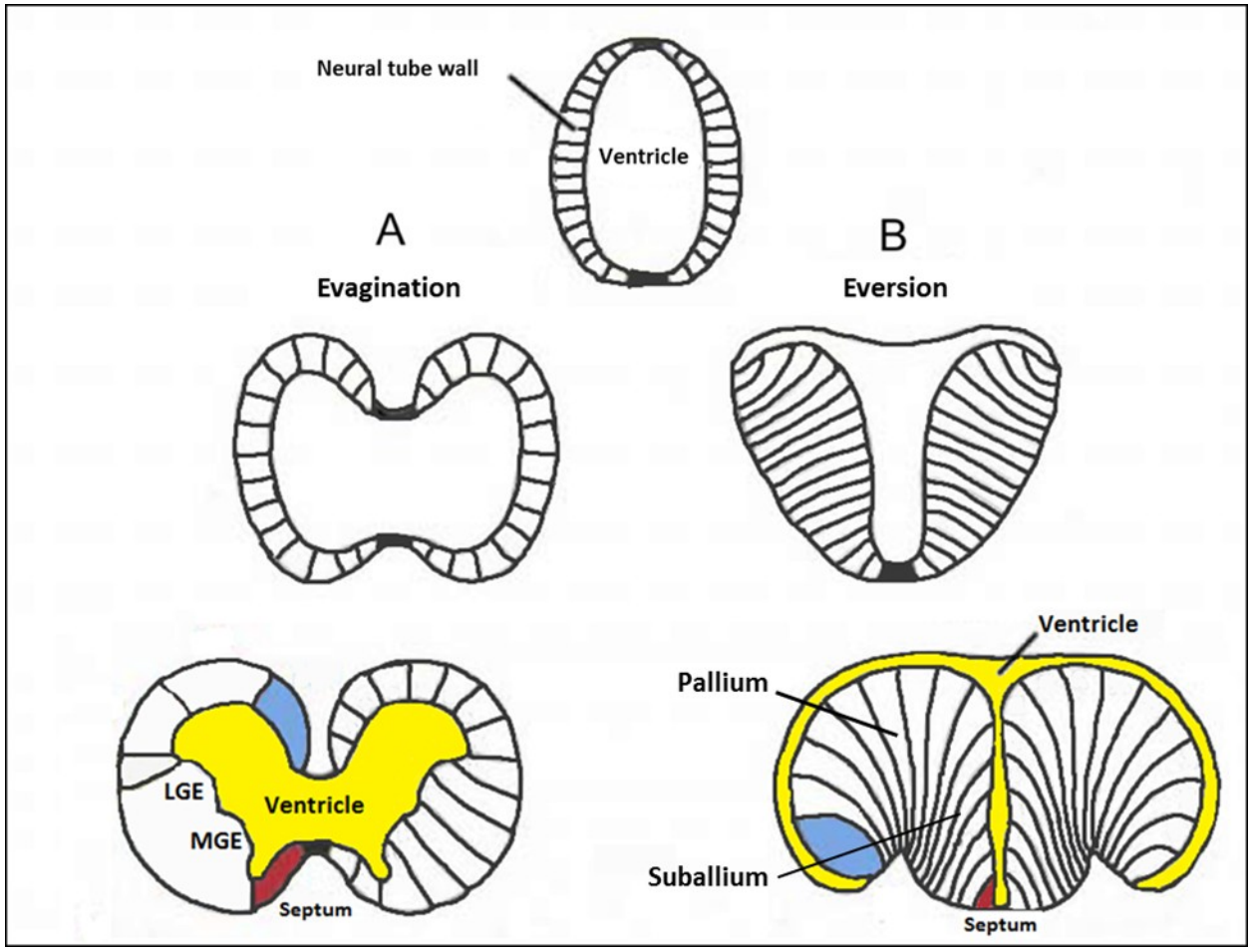

ch
ot
cb
ob
m
MAMMAL
BIRD
p
BONY FISH
REPTILE
aob
SHARK
AMPHIBIAN
LAMPREY

during development, and can be analyzed through the characterization and comparison of developmental GRN activity. To better understand how these differences in brain morphology have evolved, it is necessary to first gain a basic understanding of both the morphogenetic and molecular bases of brain development. The focus of this work is the characterization of *Dlx* transcription factor activity during the development of GABAergic interneurons in the vertebrate brain. To give context for this, the following sections will provide an overview of the transcription factors involved during early brain regionalization (1.2.3) and those implicated in neurogenesis (1.2.4).

1.2.3 Transcription factors involved in embryonic brain regionalization

The regionalization of the forebrain is dependent on the expression of a variety of signaling molecules and transcription factors. Early segmentation follows the suppression of caudalizing factors such as bone morphogenetic proteins (BMPs), fibroblast growth factors (FGFs), Wnts, and Hedgehog (Hh) (Kudoh *et al.,* 2002). A pivotal player in forebrain regionalization is the "organizer", or signaling center, from which transcription factors and other signaling molecules are produced to activate GRNs that distinguish different forebrain domains. The anterior border of the neural plate (ANB) acts as an organizer by inducing neighbouring tissues to express genes unique to the telencephalon. The ANB releases a secreted frizzled related protein (sFRP) called Tlc in mice, which inhibits the activity of Wnt caudalizing factors expression in the forebrain (Houart *et al*., 2002). Zebrafish develop without a telencephalon when there are mutations in their Wnt repressor genes *Axin1* or *Tcf3* (Kim *et al.*, 2000;

Figure 6: Cartoon contrasting anterior neural tube folding in vertebrates: evagination versus eversion. Transverse sections of the prosencephalon, dorsal is up. Following neurulation (top), neural tube walls surround a ventricle in all extant vertebrates. (A) During evagination (occurring in in all surveyed tetrapods), the neural tube squeezes medially along the dorsoventral axis to form two laterally connected ventricles surrounded larger neural tube walls. (B) During eversion (occurring in teleost fish), the neural tube wall separates on the dorsal side and grows into two large lobes, as the ventricle is compressed medially and expands dorsolaterally around the neural tube walls. Yellow regions indicate ventricles. Blue and red regions represent brain structures arising from equivalent locations of the early neural tube, but have relatively different locations after evagination / eversion. LGE- lateral ganglionic eminence, MGE – medial ganglionic eminence.


Neural tube wall
Ventricle
A
Evagination
B
Eversion
LGE
MGE
Ventricle
Septum
Pallium
Ventricle
Suballium
Septum

Heisenberg *et al.*, 2001). Conversely, high Wnt activity leads to the expression of diencephalic markers in the posterior forebrain (Braun *et al*., 2003). The ANB organizer therefore releases Wnt antagonists in a decreasing rostrocaudal concentration gradient in the forebrain. This gives rise to telencephalic gene expression where Wnt antagonist concentration is high, and diencephalic gene expression where concentration is low. It was found that the expression of *irx1b* and *otxl1/2* lead to the formation of the ZLI in zebrafish, which acts to define the boundary between the telencephalic and diencephalic forebrain regions (Schlopp *et al*., 2007). In both mice and zebrafish, this organizer releases hedgehog (Hh) proteins that activate *fgfs* underlying prethalamus and thalamus growth (Schlopp *et al*, 2006). When the expression of *wnt* is experimentally reduced in the diencephalon, the ZLI is expanded and the prethalamus is lost (Jeong *et al.*, 2007). The overexpression of *fez1* can rescue this phenotype, indicating that it too is involved in proper diencephalic development.

1.2.4 Transcription factors involved in embryonic neurogenesis

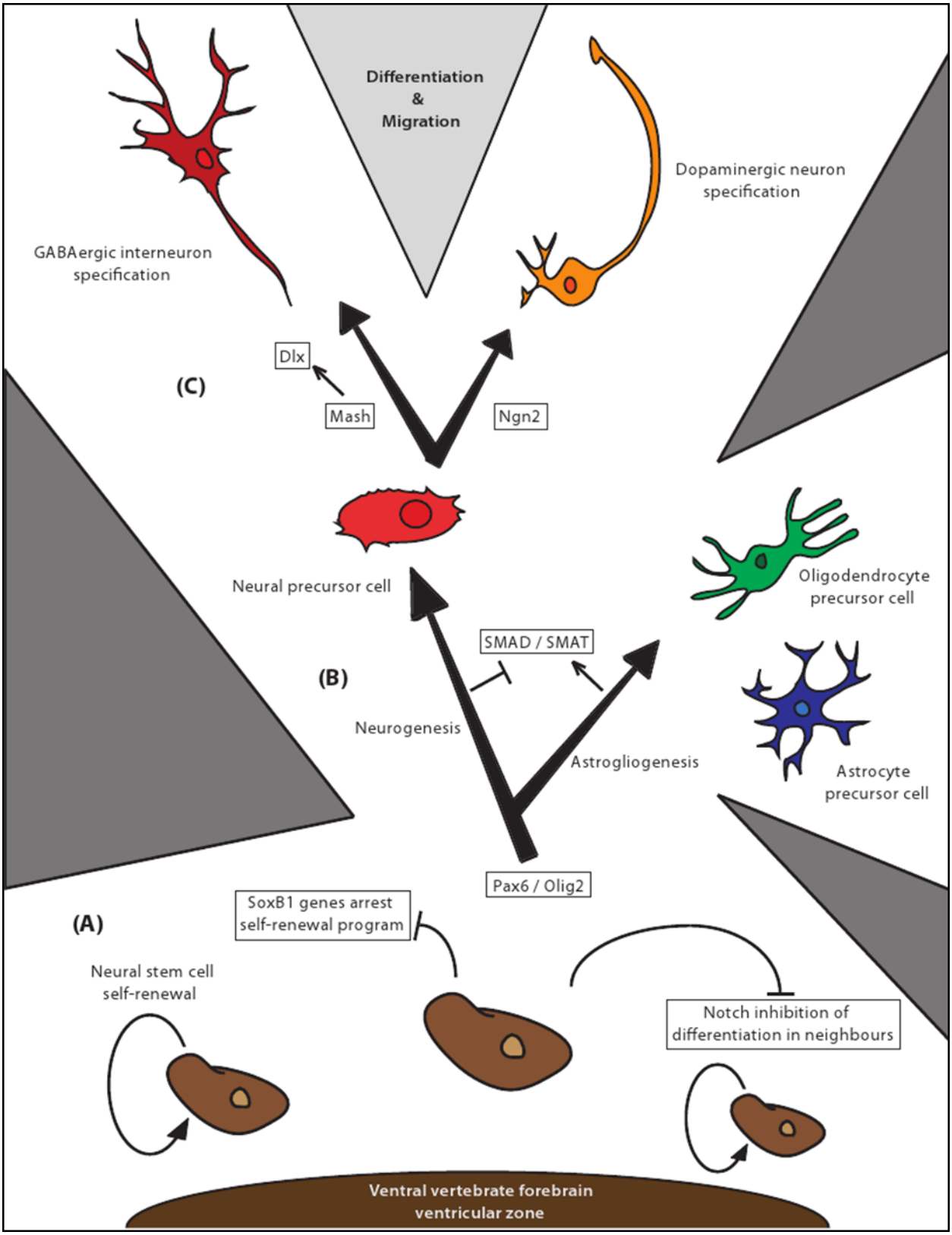
To summarize sections 1.2.1 - 1.2.3, the vertebrate neural tube undergoes folding, through either evagination or eversion, and is pinched into forebrain, midbrain and hindbrain regions. The forebrain further separates to form the telencephalon (anterior) and diencephalon (posterior). The brain is an interesting model for comparative evolutionary study because it has both highly conserved organizational features between lineages, as well as extremely derived lineage –specific novelties. The regionalization of the brain is controlled by complex cascades of gene regulatory networks that are co-ordinated by 'organizers' such as the ZLI.

Simultaneously to the regionalization of the forebrain, certain cells within the neural ectoderm undergo neurogenesis, becoming specified to neural progenitor cell (NPC) fate. The

central nervous system of vertebrates is made up of neurons and glial cells, the latter being divided into oligodendrocytes and astrocytes. Neural progenitors are created from epithelial cells located within the forebrain ventricles (Rakic, 1972). The progenitor cells renew themselves by dividing asymmetrically, where one daughter cell remains a neural progenitor and the other is committed to become one of three neural cell types: neurons, astrocytes and oligodendrocytes (**Figure 7).** By releasing neurotransmitters at synapses, the electrically - excitable neurons process and transmit information (Guillemot, 2007). Subtypes of neurons are distinguished by different neurotransmitters used, the morphology of their dendritic tree and cell body, and the cells to which their axons connect. Oligodendrocytes make the sheaths that insulate and promote the long-distance signaling of axons. Astrocytes modulate synaptic transmission, provide metabolic support to neurons, maintain water and ionic balance, and play an important structural role in the brain (Guillemot, 2007). Neurons become specified first during development, with oligodendrocytes to follow and finally astrocytes. For a brain to function correctly, neural progenitors often must migrate to specific positions, and in the central nervous system this may be done though radial or tangential migration paths. Radially migrating cells travel from an interior progenitor zone to the outer layers of the brain, whereas tangentially migrating cells travel orthogonally through the brain (Ayala *et al*., 2007).

Broadly, there are several groups of transcription factors involved in gene regulatory networks that underlie early neuron development and differentiation: Patterning, progenitor/proneural, and neuronal proteins. Patterning transcription factors subdivide the neural tube into distinct regions, initially along dorsal-ventral (DV) and anterior/posterior axes, in which cells acquire distinct positional identities. For example, in mice the homeodomain

Figure 7: Cartoon depicting the highly simplified molecular pathways underlying neurogenesis in the vertebrate ventral forebrain. (A) In the ventral forebrain ventricular zone (brown region on bottom) neural stem cells (NSCs) are in a cyle of self-renewal. With the activation of *SoxB1* family genes in certain NSCs, this cycle is interrupted and other genes (not shown) are expressed to initiate notch - mediated signalling in neighbouring NSCs (inhibiting the latter from differentiating). (B) These cells begin to express *Pax6* and *Olig2* and separate into one of two lineages undergoing either astrogliogenesis or neurogenesis. The expression of SMAD/SMAT proteins occurs during astrogliogenesis, while the inhibition of these genes occurs during neurogenesis. (C) Neural precursor cells in this region predominantly develop into GABAergic interneurons or dopaminergic neurons. The expression of the bHLH transcription factor Mash and subsequent activation of *Dlx* genes contributes to GABAergic interneuron fate specification, while the expression of Ngn2 contributes to dopaminergic fate specification (produced by author).


Differentiation
&
Migration
Dopaminergic neuron
specification
GABAergic interneuron
specification
(C)
Dlx
Mash
Ngn2
Neural precursor cell
Oligodendrocyte
precursor cell
(B)
SMAD / SMAT
Neurogenesis
Astrogliogenesis
Astrocyte
precursor cell
Pax6 / Olig2
(A)
SoxB1 genes arrest
self-renewal program
Neural stem cell
self-renewal
Notch inhibition of
differentiation in neighbours
Ventral vertebrate forebrain
ventricular zone

proteins *Pax6*, *Nkx2.2* and basic helix-loop-helix (bHLH) transcription factor Olig2 determine distinct DV domains in the developing neural tube of vertebrates (Guillemot, 2007). A key point in early neural development is that there is a coupling between spatial patterning and fate specification of cells. A cell's position along the A/P and D/V axes is reflected in the combination of transcription factors present to determine its fate. The expression of progenitor and proneural proteins leads to neuronal commitment, cell cycle exit and differentiation (Guillemot, 2007). This also initiates Notch signaling in adjacent progenitors, inhibiting their differentiation into neurons. In mice, the principle proneural proteins are *Lhx3* (LIM-homeodomain transcription factor), neurogenin (*Ngn*) *1-3*, *Math1* and *Mash1* (both basic helix-loop-helix transcription factors). *Ngn1* and *Ngn2* (also bHLH transcription factors), are expressed in the dorsal telencephalon and specify dorsal progenitors, whereas *Mash* (mouse achaete-schute homolog), is expressed in the ventral telencephalon and specifies ventral progenitors (Osorio *et al*., 2010). Neuronal proteins are expressed in post-mitotic proneural cells, and contribute to subtype-specific differentiation. These proteins include *Hb9* (DNA-binding protein), *Mbh1* (myc basic motif homolog-1), and *Dlx1/2*. The expression of proneural Ngn2 leads to a cell lineage which will later express neuronal Hb9 and eventually differentiate to become motor neurons. Similarly, cells which express proneural *Math1* will later express neural *Mbh1* and become spinal commissural neurons. Finally, and most important to my work, the expression of proneural *Mash1* will lead to the expression of neural *Dlx1/2, Dlx5/6* and contribute to the specification of GABAergic interneurons. The functional roles of GABAergic interneurons and the GRNs underlying their development will be addressed in the following two chapters.

1.2.5 GABAergic interneuron function and development

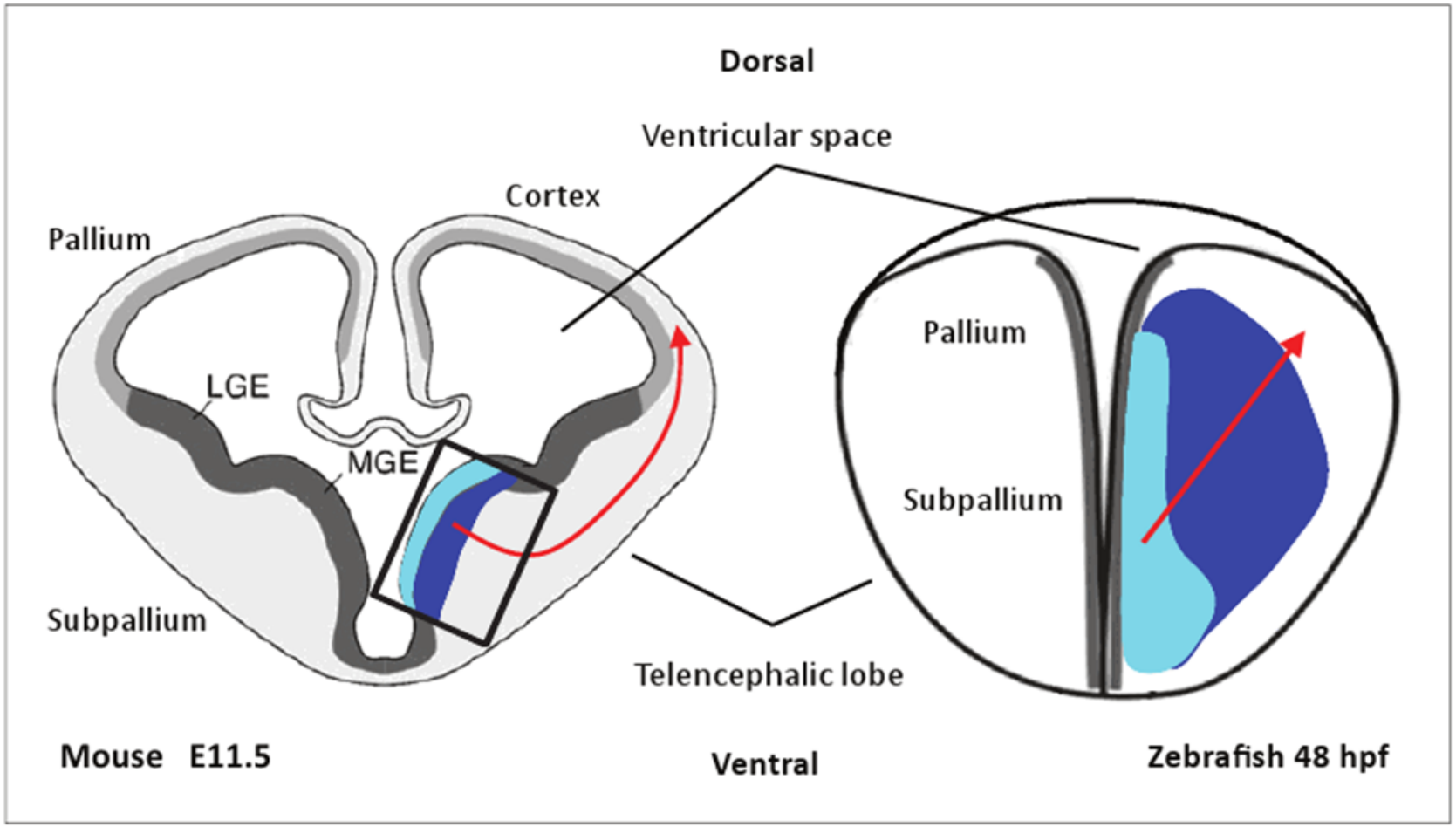
GABAergic interneurons are involved in both inhibitory and disinhibitory control of local neural circuits in the cerebral cortex, striatum and hippocampus. They make up approximately 20% of all cortical and hippocampal neurons, and 95% of all striatal neurons (Wonders and Anderson, 2006). Several subsets of GABAergic interneurons are identified based on different molecular markers including parvalbumin, somatostatin, calbindin and calretinin. Some of their roles include modulating synaptic plasticity and neuronal activity, integrating sensory information, discriminative information processing and the generation of oscillatory rhythms (Benes and Berretta, 2001). Human neurological disorders such as Autism, Epilepsy and Schizophrenia have been linked to anomalies in GABAergic interneuron function. They release γ-aminobutyric acid (GABA), which is synthesized by glutamic acid decarboxylases (*Gad*65 and *Gad67* in mice*)*. There are two known *Gad* genes in the zebrafish, *gad1a* and *gad2,* and possibly a third *gad1b* whose expression is not yet characterized (Martin *et al.,* 1998). The inhibitory action of GABA is produced by its binding to specific axon or dendrite membrane receptors and inducing the target neuron to open ion channels, causing a depolarization of action potential within the neuron. During forebrain development GABAergic interneurons arise through differentiation and migration from proliferative zones in the brain (Anderson *et al,* 1997).

1.2.6 GRNs controlling GABAergic interneuron development in mice and zebrafish

During mouse development, these neurons are first found in the telencephalic subventricular and ventricular zones of the medial and lateral ganglionic eminences, respectively, through which they tangentially migrate to finally reach the olfactory bulb, cerebral cortex, hippocampus and piriform cortex (Panganiban and Rubenstein, 2002). Similarly,

zebrafish GABAergic interneurons are born at 24 hours post fertilization (hpf), near the medial subpallial ventricular wall and migrate dorsolaterally to populate the cortex. As the teleost ventricular wall and the mammalian ventricular zones are thought to be homologous (despite the anatomical differences resulting from telencephalic eversion vs. evagination), the migration of GABAergic neurons from these regions is considered evolutionarily conserved among vertebrates. A phylotypic stage exists at 12.5-13.5 embryonic days in mice, and 2-3 days post fertilization (dpf) in zebrafish, at which time GABA positive cells in both species have extremely similar distribution patterns (Mueller *et al*., 2006). In both zebrafish and mouse *Hh* signaling in the forebrain is upstream of *Fgfs*, which positively regulate *ascl1a* and *Mash*, respectively. The proneural mouse achaete-scute homolog (*Mash*) is involved in a cascade specifying GABAergic interneuron fate in neural progenitors. At this phylotypic stage the zebrafish homolog *ascl1* (or *zash*), shows similar forebrain expression to *Mash* in mice. Specifically, they are both expressed in the subpallium, preoptic region, prethalamus and hypothalamus (Wullimann and Mueller, 2002). In the mouse ventral telencephalon, *Mash* plays a role in specifying both oligodendrocytes and GABAergic interneurons through the activation or suppression of *Dlx1*/*Dlx2*. *Dlx* gene expression co-localizes with, and is implicated in, GABAergic neuron migration and differentiation in the mouse forebrain. Similarly in zebrafish, *dlx* genes have partially overlapping expression patterns with *gad1a* during forebrain development (Macdonald *et al,* 2010). *Dlx1* and *Dlx2* are both expressed in the subventricular and ventricular zones (SVZ and VZ) of the medial and lateral ganglionic eminences, respectively (**Figure 8)**. *Dlx5* and *Dlx6* are expressed later than *Dlx1/Dlx2,* in the more-differentiated neurons found migrating in the subventricular and mantle zones (Panganiban and Rubenstein, 2002). Additionally, *Dlx1*$^{-}$*/Dlx2*$^{-}$ mice with targeted deletions of *Dlx1* and *Dlx2* have decreased GAD expression and exhibit

Figure 8: Comparative diagram of *Dlx* expression in transverse sections of the mouse (left) and zebrafish (right) telencephalon. Dark gray regions are the proliferative ventricular zones, light gray region (in mouse) is the pallial ventricular zone, light blue regions indicate the expression pattern of *Dlx1/Dlx2* (*dlx1a/dlx2a),* dark blue regions indicate expression pattern of *Dlx5/Dlx6* (*dlx5a/dlx6a*), and the red arrow signifies the migration path of GABAergic interneuron precursor cells born in the ventricular zone. LGE – lateral ganglionic eminence, MGE – medial ganglionic eminence (adapted from Wulliman *et al*, 2009).


Dorsal
Ventricular space
Cortex
Pallium
LGE
MGE
Subpallium
Telencephalic lobe
Pallium
Subpallium
Mouse E11.5
Ventral
Zebrafish 48 hpf

abnormal neurite morphogenesis in the telencephalon (Marin et al., 2000). These mice also lack correct migration of cortical interneurons from the subcortical telencephalon (Anderson et al, 1997).

The Dlx1/Dlx2 proteins negatively regulate the expression of *Olig2*, a gene involved in oligodendrocyte precursor cell (OPC) formation (Petryniak *et al.*, 2007). Instead, progenitors expressing *Dlx1/Dlx2* acquire a GABAergic interneuron fate. Mash acts here as a switch: when it activates *Dlx1/Dlx2* expression in a given cell, *Olig2* is suppressed and an interneuron fate is specified, whereas when it does not activate *Dlx, Olig2* is expressed and progenitors become OPCs. Necdin is a member of the MAGE (melanoma antigen protein) family that promotes neural differentiation, suppresses cell proliferation and inhibits death of certain cell lineages (Kuwajima *et al.*, 2006). By forming a complex with an intermediary MAGE-D1 protein of the same family, necdin binds the Dlx2 and Dlx5 proteins, contributing to GABAergic fate specification. *Arista-less homeobox* (*Arx*) has an upstream enhancer strictly bound by Dlx2 in the mouse ventral forebrain and mediates the timing of GABAergic interneuron progenitor migration (Colasante *et al.,* 2008). Furthermore, necdin has been shown to greatly increase the affinity of Dlx2 binding to the *Arx* enhancer.

The preoptic area is another source of migrating GABAergic interneurons to the cortex, and it is located in the diencephalon (Gelman *et al.*, 2009). GABA-positive cells have also been found migrating radially and tangentially in the thalamus (diencephalon) (Ortino *et al*., 2003). They originate in a region of the prethalamus called the reticular nucleus, and migrate to the dorsal thalamus. This migration is at least partially induced by the activity of homeoprotein Otx2 (Inverardi *et al.,* 2007).

In zebrafish, ventral forebrain expression of *dlx2a* is lost following morpholino-mediated knockdown of *fgf3* and *fgf8* (Miyake *et al*., 2005). In the same morphants *gad1a* and *ascl1a* are also reduced, but only in the prethalamus. When *fgf19* is knocked down, however, *gad1a* and *gad2* expression is lost in the telencephalon and prethalamus (Miyake *et al*., 2005). *Ascl1a* is also positively regulated by *her6* (by decreasing *neurog1* repression), which, along with *fgf*s, is activated by *Hh* (Scholpp et al. 2009). In the zebrafish diencephalon, there are two parallel cascades that are thought to regulate the formation of GABAergic inteurons. Along with the above *Hh*-mediated cascade, there is a *fez1*-mediated cascade that contributes to *ascl1a, dlx* and *gad1* expression (Braun *et al.,* 2003). Fez1 is a repressor of *wnt* caudalizing factors but an activator of *wnt8b* and *lef1*, both of which contribute to the expression of *ascl1a* in progenitors and *dlx2a* in post-mitotic neurons. It is clear that the zebrafish forebrain has regionally unique GRNs underlying GABAergic interneuron development. The regulatory relationships between zebrafish *ascl1a, dlx1a/dlx2a, dlx5a/dlx6a* and *gads* are not as well characterized as in mice. Expression patterns of orthologous genes in mice and zebrafish are highly similar, but it is not known if they exhibit similar regulatory relationships. The following section will give an overview of the function and diverse developmental roles played by *Dlx* genes in vertebrates, with a focus on the forebrain.

1.2.7 The role of *Dlx* genes in vertebrate development

Homeodomain transcription factors are encoded by a highly conserved group of genes, the homeobox genes, which are present in all metazoans and play a crucial regulatory role in early embryonic development. One family of homeobox genes, known as *distal-less homeobox (Dlx)* genes, is present in all surveyed vertebrate species including human. A highly conserved

homeodomain (DNA-binding) of 61 amino acids is shared between all *Dlx* genes (Liu *et al*., 1997). DLX proteins recognize *in vitro* a consensus binding site of (A/C/G) TAATT (G/A) (C/G) (Feledy *et al*., 1999). They have the capacity to form homodimers and heterodimers with other factors to initiate transcription, but may sometimes function as repressors (Zhang *et al*., 1997; Masuda *et al*., 2001; Le *et al.,* 2007). These genes are expressed during the development of the forebrain, branchial arches and derivatives, sensory organs and limb buds. Depending on the species, the *Dlx* gene family consists of six to eight genes, mainly organized in convergently transcribed bigene pairs. These genes are orthologous to the *Distal-less* (*Dll*) gene found in *Drosophila melanogaster*, which plays an important role in proximodistal patterning in early limb and antennae formation, and also contributes to sense organ development (Cohen and Jurgens, 1989; Cohen *et al*., 1989). The last common ancestor between *Drosophila* and vertebrates had one copy of this gene, and it is likely that a tandem gene duplication event occurred to create a convergently transcribed gene pair in the lineage leading to vertebrates. Right before vertebrate diversification there were two known genome duplication events, which can account for the presence of up to three gene pairs found in some vertebrate species today. Each *Dlx* pair is linked to unique Hox gene clusters, which are key regulators of body segmentation. Divergence in *Dlx* paralog function following a gene or genome duplication event was likely an important mechanism contributing to the diversification of their developmental roles (Ghanem *et al*., 2003).

During mouse telencephalon development, newborn cells in the proliferative region of the ventricular zone express *Dlx1* and *Dlx2*. Cells in the subventricular zone begin expressing *Dlx5,* and finally *Dlx6* in differentiating and most-mitotic neurons in the more lateral mantle zone (Liu *et al.*, 1997). The expression of *Dlx* genes shows a considerable overlap with those of

Gad genes in the forebrain. In *Dlx1*$^{-}$*/Dlx2*$^{-}$ mice with targeted deletions of *Dlx1* and *Dlx2,* there is an accumulation of neural progenitors in the ventricular proliferative zone (Anderson *et al*., 1997a). This is thought to be due to an arrest of neurogenesis, and the prevention of proper progenitor differentiation and migration. As a result, the number of GABAergic interneurons that tangentially migrate to the cerebral cortex and olfactory bulb is significantly decreased (Anderson *et al.*, 1999; Bulfone *et al.,* 1998; Anderson *et al.,* 2001). It is essential to note that the development of dopaminergic and cholinergic neurons is also negatively affected. In these same mutants, *Dlx5* and *Dlx6* forebrain expression is drastically decreased, indicating that *Dlx1/Dlx2* are essential in their regulation (Zerucha *et al.,* 2000; Stühmer *et al., 2*002a; Stühmer *et al., 2*002b). In *Dlx5/Dlx6* double null mutants, mice develop with exencephaly (failure of anterior neural tube closure), and therefore it is impossible to assess their function during GABAergic interneuron development. Because *Dlx5/Dlx6* forebrain expression is contingent on that of *Dlx1/Dlx2,* it is difficult to determine which components of the *Dlx1*$^{-}$*/Dlx2*$^{-}$ phenotype are caused by the down-regulation of *Dlx1/Dlx2* versus that of *Dlx5/Dlx6. Dlx* genes are also expressed in the zebrafish forebrain, but further investigation is required in order to gain a better understanding of their function. To better understand the role played by *Dlx* genes in vertebrate forebrain development, it is important to study the activity of *cis*-regulatory elements (CREs) acting on these genes.

1.2.8 CRE regulation of *Dlx* bigene clusters

A crucial path to a more comprehensive understanding of *Dlx* gene activity is through the study of their regulators. The regulation of gene expression is profoundly complex, involving changes in the 3-dimensional structure of DNA, the binding of transcription factors, the

recruitment of polymerases, post-transcriptional and post-translational modification, to name just a few components of this process. CREs have been found in the intergenic and upstream flanking regions of *Dlx* bigene clusters. Between all surveyed vertebrates, *Dlx* protein-coding sequences are 75% identical over several hundred base pairs (Ghanem *et al.,* 2003). Conversely, the sequences of these CNEs are relatively less conserved among lineages than the neighbouring coding sequences. The widespread presence of homologous CNEs indicates they have some importance in regulating the neighbouring *Dlx* genes. Both genes in a given bigene cluster have overlapping expression patterns. This could be explained by the binding of transcription factors onto intergenic CREs such as enhancer sequences, to create a transcription factor-enhancer complex that interacts with the promoters of both neighbouring *Dlx* genes, driving similar expression.

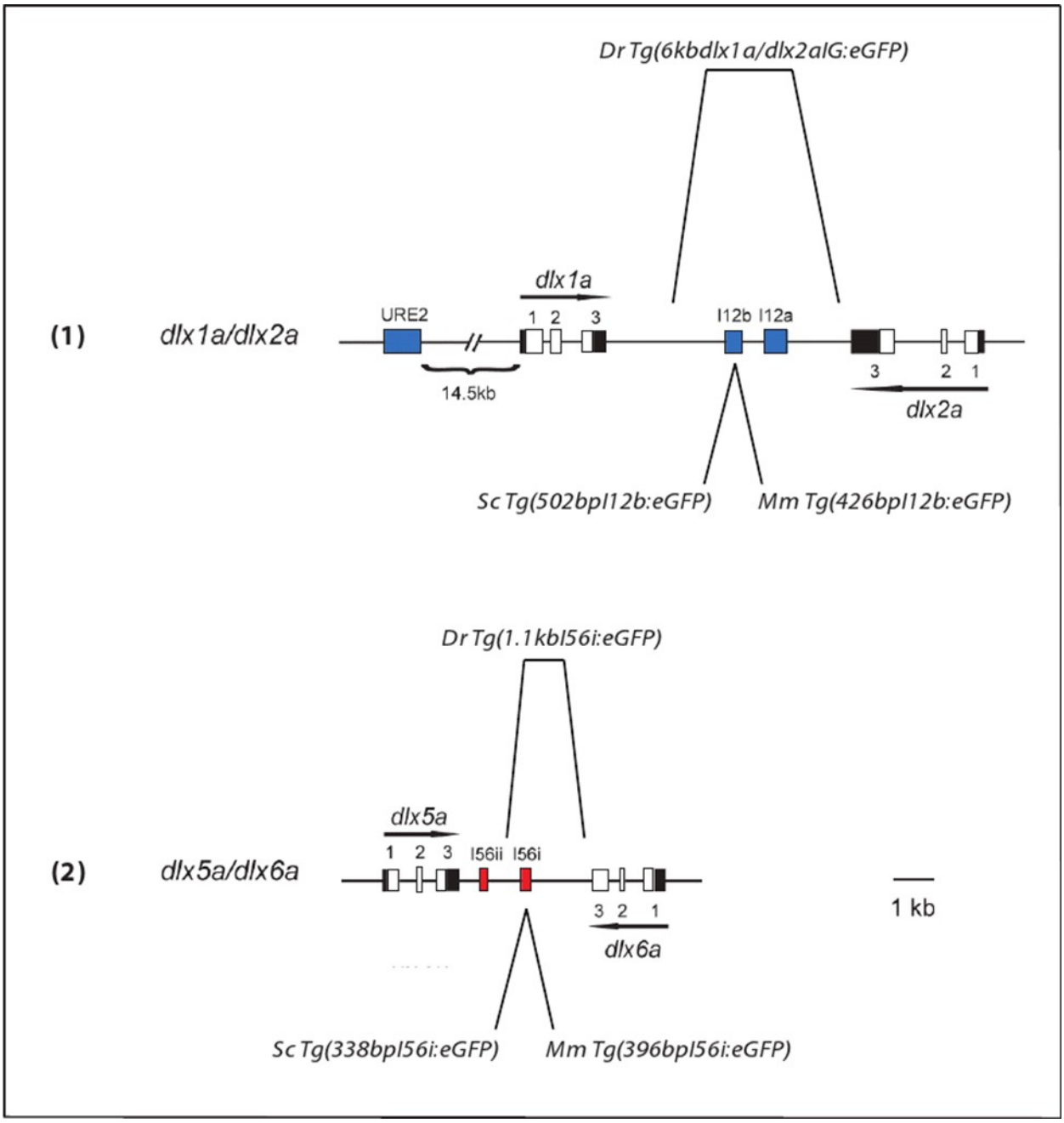
In the intergenic region between *Dlx1* and *Dlx2* (in both mice and zebrafish), there are two CNEs, I12a (550bp) and I12b (450bp) (Ghanem *et al.*, 2003). I12a and I12b are able to drive reporter gene expression in the branchial arches and the forebrain, respectively (Park *et al*., 2004). Also, I12b is known to act as an enhancer when bound by the MASH1 transcription factor, playing some role in driving *Dlx1/2* gene expression in the brain (Poitras *et al,* 2007). Furthermore, the Dlx2 protein can itself bind to I12b, possibly causing a positive-feedback in expression. The intergenic region of the *Dlx5/6* bigene cluster contains two putative enhancers, I56i (400 bp) and I56ii (300 bp). Between mice and zebrafish, orthologous enhancers share approximately 85% sequence similarity (Zerucha *et al*., 2000). In zebrafish and mice, a 1.4 kb intergenic sequence between zebrafish *dlx5a* and *dlx6a* is sufficient to drive reporter gene expression identically to that of *dlx5a/dlx6a* (*Dlx5/Dlx6)* in the forebrain (Zerucha *et al*., 2000). It has also been shown that the Dlx2 protein interacts with I56i, to drive reporter gene expression

that coincides with *Dlx5* and *Dlx6* expression patterns. This suggests a cross-regulatory cascade among *Dlx* genes (Stühmer *et al.,* 2002). The regulatory functions of these enhancers have been conserved since the divergence of mouse and zebrafish lineages. Interestingly, although I12b, I56i and I56ii all direct expression to the forebrain, there is little sequence similarity between them. Two other enhancers were found upstream of *Dlx1*(*dlx1a*), and named Upstream Regulatory Elements 1 (URE1) and 2 (URE2) (Thomas *et al.*, 2000). URE2 can target reporter expression to the forebrain in a similar fashion to the above three enhancers (**Figure 9)**. Conversely, URE1 only drives reporter expression in the retina, a tissue in which *Dlx1* and *Dlx2* are later expressed (Dollé *et al*., 1992). Mouse URE2, I12b and I56i enhancers activate reporter genes in interneuron progenitors of the ganglionic eminence (Potter *et al*., 2009). Conversely, I56ii primarily marks post-mitotic neurons (Ghanem *et al.*, 2008).

The rest of the intergenic sequences (non-conserved) have not been tested for putative enhancer activity. In the cases of I12b and I56i, it is possible that enhancing activity is contingent on their specific location within the intergenic sequence. Another interesting contributor to the complex regulation of *Dlx* genes is *Evf2*, a long non-coding RNA which is transcribed from the intergenic region of *Dlx5/Dlx6* (Feng *et al.,* 2006). Its role is to help recruit Dlx2 to the *Dlx5/Dlx6* intergenic enhancers, probably by forming a DNA-RNA-protein complex (Bond *et al.,* 2009).

To better identify whether sequence differences in homologous *Dlx* intergenic enhancers contribute to differential *Dlx* expression during development, a comparative look can be taken between distantly related species. Mammalia, Osteichthyes and Chondrichthyes are three vertebrate classes within which early gene regulatory comparisons can be made. Mice (*Mus musculus*), zebrafish (*Danio rerio*) and dogfish (*Scyliorhinus canicula*) are representatives of

Figure 9: Schematic of genomic organization of *dlx* bigene clusters in the zebrafish. 1) *Dlx1a/dlx2a* convergently transcribed genes have CREs (blue boxes) located in the intergenic region (I12b and I12a), as well as 14.5 kb upstream of *dlx1a* (URE2). A preexisting transgenic line, Dr Tg(6kbdlx1a/dlx2aIG:eGFP), has a 6 kb intergenic fragment including the zebrafish (*Danio rerio*) I12b enhancer upstream of a beta-globin minimal promoter and eGFP reporter gene. For comparative analysis of orthologous enhancer activity of dogfish and mouse species in the zebrafish genetic background, a 502 bp I12b sequence was amplified from the dogfish (*Scyliorhinus canicula*) genome, Sc Tg(502bpI12b:eGFP), and a 426 bp I12b sequence was amplified from the mouse (*Mus musculus*) genome, Mm Tg(426bpI12b:eGFP). Fragments were bidirectionally subcloned into pSP72 plasmid, which contains a beta-globin minimal promoter upstream of eGFP. 2) *Dlx5a/dlx6a* convergently transcribed genes have CREs (red boxes) located in the intergenic region (I56ii and I56i). A preexisting transgenic line, Dr Tg(1.1kbI56i:eGFP), contains 1.1 kb zebrafish intergenic region which includes the I56i enhancer. To compare activity between zebrafish, mouse and dogfish orthologous enhancers in a zebrafish genetic background, I amplified and subcloned a 338 bp I56i fragment from the dogfish genome, Sc Tg(338bpI56i:eGFP), and a 396 bp fragment from the mouse genome, Mm Tg(396bpI56i:eGFP) bidirectionally into pSP72 plasmid. The numbered white boxes are coding regions, and black boxes are untranslated regions of exons.


Dr Tg(6kbdlx1a/dlx2aIG:eGFP)
(1)
dlx1a/dlx2a
dlx1a
URE2
1 2 3
I12b I12a
3 2 1
14.5kb
dlx2a
Sc Tg(502bpI12b:eGFP)
Mm Tg(426bpI12b:eGFP)
Dr Tg(1.1kbI56i:eGFP)
dlx5a
(2)
dlx5a/dlx6a
1 2 3
I56ii I56i
3 2 1
dlx6a
1 kb
Sc Tg(338bpI56i:eGFP)
Mm Tg(396bpI56i:eGFP)

these classes. It is known that *Dlx* expression patterns differ slightly between these three vertebrate lineages, which have evolved distinct brain morphologies. As an example, in both mouse (Zerucha et al, 2000) and zebrafish (MacDonald *et al.*, 2010) the *Dlx1/Dlx2 – dlx1a/dlx2a* (zebrafish) tandem is expressed earlier than the *Dlx5/Dlx6 – dlx5a/dlx6a* tandem in the subventricular zone of the telencephalon while, in the dogfish, it is the *Dlx5/Dlx6* tandem that is expressed earlier than the *Dlx1/Dlx2* tandem in the orthologous zone (personal communication from Melanie Thibaud-Debiais). Lineage specific changes in how *Dlx* genes are regulated may have given rise to the above temporal variations in expression. The divergence in orthologous CRE sequences could mean they are bound by different transcription factors, or there are different binding affinities for certain transcription factors. This could (1) change the geospatial expression of *Dlx* genes during development, (2) result in altered migration and differentiation of GABAergic interneurons and therefore (3) alter brain phenotype of mature individuals. This phenotype could then be subject to natural selection and contribute to the varied morphological forms seen in vertebrate brains. What is less clearly known is whether the change in *Dlx* expression can in fact account for some of the seen morphological differences between vertebrate species.

1.3 Statement of purpose

Not only do *Dlx* genes have deep historical roots in vertebrate evolution, they are also central in embryonic development, making the family an ideal subject for the study of evo-devo. As evo-devo makes cross-species comparisons of the regulatory interactions between genes during development, such an approach has been taken in this work. The *Dlx* homeodomain transcription factor gene family is highly conserved in vertebrates and plays various important

roles during development. In mice, several *Dlx* genes are necessary for the correct migration and differentiation of GABAergic interneurons in the forebrain. It is of interest to determine whether orthologous genes have a similar function in the developing zebrafish brain, and to what degree the GRNs underlying GABAergic interneuron formation are conserved between the two lineages.

One approach is to compare the regulatory activity of orthologous enhancer sequences. For the first chapter of my thesis I intended to produce transgenic zebrafish lines with reporter genes driven by individual enhancers from mice, dogfish and zebrafish *Dlx* intergenic regions. As I12b and I56i are activated in this GRN in mice and zebrafish, they were good candidates for cross-species comparison. If orthologous enhancers from mice and dogfish show unique reporter gene expression patterns when inserted into the zebrafish (a heterologous background), compared to those of zebrafish (a homologous genetic background), this would signify that enhancer sequence conservation does not necessarily reflect functional conservation between distant species. I was unable to establish stable transgenic lines, but instead looked at transient expression of enhancer-reporter gene constructs in zebrafish larvae. The I12b and I56i enhancers from all three species are able to drive eGFP reporter expression in forebrain regions, and consistent to where *dlx1a/dlx2a* and *dlx5a/dlx6a* are expressed endogenously. This suggests the functional roles of orthologous enhancer sequences have not undergone significant divergence.

The second chapter of my thesis sought to identify the role played by zebrafish *dlx* genes in forebrain developmental GRNs, specifically elucidating whether the *MASH* ortholog *ascl1a* is an activator of *dlx* expression, and if this contributes to the migration and differentiation of GABAergic interneurons. Do GRNs involving *dlx* genes in the developing forebrain have a conserved architecture and function among vertebrate species? What are the downstream effects

if *dlx* genes are experimentally down -regulated? Can this phenotype be rescued by exogenous expression of intermediary genes in the GRN? My work has helped identify interesting similarities and differences between the mouse and zebrafish GRNs involved in forebrain development in the following three ways: (1) Down-regulation of *ascl1a, dlx1a/dlx2a* or *dlx5a/dlx6a* does not have a visible effect on the abundance of GABA-producing neurons (*gad1a* expression) in the developing telencephalon of zebrafish, (2) down-regulation of *ascl1a* or *dlx1a/2a,* but not *dlx5a/dlx6a,* reduces the number of GABA-producing cells in the prethalamus (diencephalon), and (3) the exogenous expression of *dlx* genes can partially rescue the diencephalic *gad1a* phenotype seen when *ascl1a* is down-regulated.

2. MATERIALS AND METHODS

2.1 Amplification of enhancer fragments from BACs

Primers were designed to amplify zebrafish and dogfish I12b and I56i enhancer sequences from bacterial artificial chromosomes (BACs) containing the respective *dlx* gene pairs. **See Table 1 for** the primer sequences. Plasmids containing mouse enhancers were already available. Primers were resuspended to a concentration of 0.1 mM. Each 50 µL PCR sample contained BAC template DNA (5-10 ng per reaction), 2 µL each of left and right primer, 0.5 mM dNTPs, 5 µL of 10X transcription buffer (500 mM KCl, 100 mM Tris-HCl, pH 9.0 at room temp), 5 units of *Taq* polymerase (5 units/µL), and sterile H_20. The PCR reactions consisted of 31 cycles of a 1 min denaturation step at 94°C, a 1 min annealing step at 55°C (variable, around 5°C lower than primer Tm), and a 2 min elongation step. A volume of 1 µL of each sample was mixed with 1 µL of 10X loading dye, 8 µL of H_20 and run on a 1% agarose electrophoresis gel to verify success of PCR reactions.

2.2 Ligation of fragments into pDrive vector

Amplified DNA fragments were immediately ligated into pDrive vectors by adding 2 µL of PCR product to 2 µL of sterile H_20, 5 µL of 2X mix, 1 µL of pDrive vector (Qiagen kit, cat.# 231124), and kept at 14°C overnight. For transformation of the plasmid, 100 µL of competent E. *coli* bacterial cells were thawed on ice for 30 min, and 10 µL of ligation reaction was added. The bacteria were left on ice for another 20 min and then heat-shocked at 42°C for 1 min 15 sec. 1 mL of LB media (500mL stock = 5g tryptone, 2.5g yeast extract, 5g NaCl, H_20) was subsequently added to bacteria and left to incubate at 37°C for 1 hr. A volume of 35 µL of X-

galactosidase and 10 µL of IPTG was spread upon an LB-agar culture plate containing kanamycin. Bacteria were centrifuged at 8000 g for 1 min and 900 µL of LB liquid was removed, and the remaining 100 µL of bacteria was spread on the culture plate. The plate was incubated at 37°C overnight. Ten white colonies were selected from each plate and inoculated into 4 mL cultures of LB containing 0.4 µL of ampicillin, and left to shake at 37°C overnight. To extract plasmid DNA from each 4 mL cultures, 2 mL were transferred to a 2 mL microtube and centrifuged at 19000g for 3.5 min. All liquid was removed from the tube and 200 µL of buffer P1 (50 mM Tris HCl pH 8.0, 10 mM EDTA and 100 µg/ml RNASE A) were added, the pellet was then vortexed to resuspend bacteria. Cells were lysed by adding 200 µL of buffer P2 (1M NaOH, 10% SDS and H_20) and inverted 6 times. To separate the DNA from other cellular material, 200 µL of buffer P3 (75% EtOH, 25mM NaCl, 5mM Tris-HCl, pH7.5) was immediately added, the tube was inverted 6 times and centrifuged at 18000 g for 10 min. Supernatant was transferred to a new 1.5 mL tube and 800 µL of 100% EtOH was added. DNA was allowed to precipitate at -20°C for ½ hr. Tubes were centrifuged at 18000 g, 4°C for 15 min to pellet the plasmid DNA. The liquid was removed from tube and replaced with 800 µL of cold 70% EtOH. The tube was centrifuged again at 18000 g, 4°C for 10 min. The ethanol was removed and the pellet allowed to air dry. DNA was eluted into 20 µL of sterile H_20, 10 µL of which was digested (with 1 µL *Sal*1 restriction endonuclease, 2 µL of 10X reaction buffer and 7 µL H_20) and run on a gel to verify that the desired fragment was inserted in the pDrive vector. A PCR test was also done to verify this. One of the 10 cultures was chosen to inoculate 50 mL of LB (containing 5 µL of ampicillin) and left shaking overnight at 37°C. A volume of 1 mL of liquid culture was removed to keep as a stock solution. This was mixed with 1 mL of 100% glycerol in a 2 mL microtube and stored at -80°C. The remaining liquid culture was pelleted at

6000 g for 15 min at 4°C and the DNA extraction protocol for the HiSpeed Midi kit was followed (Qiagen cat. #12643). Inserts were subsequently sequenced to verify correct amplification.

2.3 Enhancers subcloning from pDrive into the pSP72 vector

A total of 5 µg of pDrive containing an enhancer insert was digested using 1 µL *Sal*I restriction endonuclease 5 µL of 10X digestion buffer and H_20 to a volume of 50 µL, at 37°C for 2 hrs. A volume of 5 µL of 10X loading dye was added to the tube and the entire sample was run on a 1% agarose gel. The enhancer sequence was extracted and purified from the gel according to the Qiagen Gel Purification kit, (cat. #28706). One microgram of pSP72 vector (Promega cat. #P2191) containing *eGFP* and two Tol2 transposase recognition sites was linearized using *Sal*I restriction endonuclease. The purified enhancer fragment was ligated into pSP72 vector. The reaction solution contained 4 µL 5X ligation buffer, 1 µL T4 ligase (Invitrogen cat. #15224017), approximately 30 ng of linearized vector, approximately 30 ng of insert and H_20 to a volume of 20 µL. The reaction was kept at 14°C overnight. A volume of 16 µL of this solution was transformed into competent E.*coli* cells as described above, and grown on LB/Agarose culture plates containing ampicillin. Ten colonies were grown in liquid culture and tested using PCR and restriction endonuclease digestion to verify correct insertion of enhancer. One colony was chosen to inoculate a 50 mL liquid culture, which contained ampicillin. A volume of 1 mL of liquid culture was removed to keep as a stock solution. The remaining liquid culture was pelleted at 6000 g for 15 min at 4°C and the DNA extraction protocol for the HiSpeed Midi kit was followed. Several aliquots were made at a concentration of 175 ng/µL. For each enhancer, two plasmids were produced having both forward and reverse orientations.

2.4 *Tol2*, *dlx* and probe RNA *in vitro* transcription

Tol2 is a DNA transposon sequence initially found in the medaka (*O. latipes)* genome that encodes a transposase enzyme. This enzyme recognizes specific sequences that flank the Tol2 sequence and can therefore be used to transfer a plasmid insert with these flanking regions, into genomic DNA. For zebrafish transgenesis, Tol2 RNA and plasmid DNA can be injected into embryos simultaneously in order to increase efficiency of incorporation. For the experiments in which *dlx* genes were exogenously expressed, *dlx2* and *dlx5* RNA was synthesized *in vitro. Tol2* RNA was synthesized using the mMessage mMachine SP6 kit (cat #AM1340), consisting of 1 µg of template plasmid, 8 µL of 2X NTP/CAP, 2.2 µL of 10X reaction buffer, 2 µL SP6 enzyme mix, and sterile water (total reaction volume of 22 µL). The reaction was kept at 37 °C for 3 h. The template DNA was degraded by adding 2 µL of DNAse I to the tube and incubated for 15 min at 37 °C. To precipitate the RNA, 25 µL of 5M LiCl and 75 µL of 100% EtOH were added, and the reaction was left overnight at -20 °C. *Dlx2* and *Dlx5* were synthesized with previously made plasmids, linearized with *BamHI* and transcribed with T7 RNA polymerase. Digoxigenin-11-dUTP- labelled RNA probes for *in situ* hybridization were made through *in vitro* transcription. For each probe, 1-2 µg of plasmid containing a sequence matching part of an RNA molecule of interest, was linearized and combined with 2 µL DIG RNA labelling mix (Roche, cat #11277073910), 2 µL of 10X transcription buffer, 1 µL of RNaseOUT™ Recombinant Ribonuclease Inhibitor (Invitrogen cat #10777-019), 2 µL of either SP6/T7/T3 RNA polymerase (T7, Roche cat #10881767001, SP6, cat #10810274 001, T3, cat #11031163001), and nuclease free H_2O to a final volume of 20 µL. The reaction was kept at 37 °C for 3 h. The template DNA was degraded by adding 2 µL of DNAse I to the tube and incubated again for 15 min at 37 °C.

The RNA probe was precipitated with 25 µL of 5M LiCl and 75 µL of 100% EtOH, and left overnight at -20 °C. The RNA was centrifuged at 18000 g at 4 °C for 15 min to separate the RNA pellet from the supernatant. The supernatant was discarded, 100 µL of cold 70% EtOH were added, and the tube was again centrifuged at 18000 g at 4 °C for 15 min. The liquid supernatant was subsequently discarded and the pellet left to air dry for ½ hr. Pellets were resuspended in 30 µL RNAse-free H_2O, 29 µL formamide and 1 µL RNaseOUT™ and stored at -20 °C. Because *dlx* and *Tol2* RNA were injected into live embryos, a phenol-chloroform extraction was carried out to remove all (potentially toxic) proteins present during *in vitro* transcription. To do this, 113 µL of sterile H_20 and 15 µL of ammonium acetate were added to the 22 µL reaction, making a final volume of 150 µL. A volume of 150 µL of phenol was added and the solution was vortexed briefly and spun at 18000 g for 5 min. The top phase was removed and transferred to a new microtube, into which 150 µL of chloroform were added. The tube was again vortexed briefly and spun at 18000 g for 5 min. The top phase was removed and transferred into a new tube, and 450 µL of cold 100% EtOH was added. The RNA was allowed to precipitate overnight at -20 °C. The RNA was pelleted in the same way as the probe RNA, but the pellets were resuspended into 300 µL of RNAse-free H_2O.

2.5 Microinjection of plasmids, RNA and morpholino oligonucleotides into zebrafish embryos

Zebrafish embryos were collected at the 1-cell stage, submerged in 0.5 % bleach for 2 min then kept in E3 embryo medium (2L, 60 X stock: 34.8 g NaCl, 1.6 g KCl, 5.8 g $CaCl_2$-$2H_2O$, 9.78 g $MgCl_2$-$6H_20$, water; for 1L of 1 X solution, add 100 µL of 1% methylene blue). Embryos were arranged in rows on a 1.5% agarose injection plate moulded with a plastic piece with

straight wedges. Injections were performed under a stereomicroscope at a 6 X magnification. Micropipettes were made by heating and pulling 1mm borosilicate glass capillary tubes (Sutter Instrument cat # BF100-50-10). The tubes were cut into needles of an appropriate length to prevent embryo damage (if too short), as well as to avoid being too brittle and flexible (if too long). A working solution was prepared containing 0.05% phenol red (visible marker for injecting solution into embryo), 30 ng/µL of plasmid, 44 ng/µL of *Tol2* RNA, and sterile water. The 1 nL injection volume was calibrated by placing an injected droplet of working solution into a drop of mineral oil on a micrometer slide and using the diameter of the droplet to calculate its volume. Approximately 30 pg of plasmid and 44 pg of *Tol2* RNA were injected into every embryo. Embryos were screened at 24 and 48 hpf for transient eGFP expression in the forebrain. Individuals with prominent forebrain expression were saved and raised until adulthood.

Morpholino oligonucleotides (MOs) were purchased from GeneTools and resuspended with H_2O to a concentration of 2 mM. **See Table 2** for MO sequences. Single MO injection concentration was 4 ng/µL, and two co-injected MOs each had a concentration of 2 ng/µL, resulting in about 4 pg of MO injected into each embryo. For co-injection of MOs and RNA, the concentration of the latter was 44 ng/µL. To prevent the larvae from developing pigmentation, 1 µL/mL of 1-phenyl 2-thiourea (PTU) was added to the embryonic media at 24 hpf. Larvae were dechorionated at stages between 24-48 hpf, and fixed overnight at 4 °C in 4 % paraformaldehyde (PFA). They were then dehydrated in 100% MeOH and stored at -20 °C.

2.6 Alcian blue staining of developing skeleton

To detect morpholino oligonucleotide- induced malformations in the craniofacial skeleton, alcian blue cartilage staining was performed on 5 dpf zebrafish larvae. Larvae were

briefly washed three times in phosphate-buffered saline with 0.1% tween 20 (PBST), and bleached in 30% H_20_2 for 2 hrs to remove pigment. The larvae were briefly rinsed twice in PBST and transferred into a solution of alcian blue (1% concentrated HCl, 70% EtOH, 0.1% Alcian blue) and stained overnight. They were briefly rinsed 3 times in acidic ethanol (5% concentrated HCl, 70% EtOH) and for 20 min in HCl-EtOH. Through a series of dilutions, larvae were gradually rehydrated into distilled H_2O. Craniofacial morphologies were determined using a dissecting scope.

2.7 Whole mount *in situ* hybridization

In situ hybridization was used to qualitatively visualize gene expression patterns in whole mount larvae. Using small plastic baskets in 24-well plates, larvae were rehydrated into PBST (Phosphate buffered saline, Amresco cat #E404-200TABS, with 0.1% tween 20) through 5 min washes in MeOH 75% / PBST 25%, MeOH 50% / PBST 50%, MeOH 25% / PBST 75% followed by four 5 min washes in PBST. To allow reagent penetration of tissue, 10 µg/mL of proteinase K was used to permeabilize larvae for 5 min (24 hpf larvae) and 12 min (48 hpf larvae). To remove proteinase K larvae were washed for 5 min in PBST. Larvae were re-fixed in 4 % PFA for 20 min and washed 5 times for 5 min in PBST. Larvae were transferred to microtubes and allowed to pre-hybridize in 800 µL of hybridization mix (50% formamide, 5X SSC , 92 µM citric acid pH 6, 0.1% tween and sterile water), at 70 °C for 2 hrs. For probe hybridization, this mix was removed and replaced with 100-200 ng of probe in 200 µL of hybridization mix 'plus' (containing 50 µg/mL heparine and 500 µg/mL tRNA). Larvae hybridized overnight at 70 °C. To fully remove probe and hybridization mix, larvae were gradually transferred into 0.2 X SSC (20X stock: 3 M NaCl, 300 mM Na-Citrate, sterile H_2O, pH

7) through 15 min washes at 70 °C in 75% Hyb. mix / 25% 2 X SSC, 50% Hyb. mix / 50% 2 X SSC, 25% Hyb. mix / 75% 2 X SSC, 100% 2 X SSC, followed by two 30 min washes in 0.2 X SSC. The larvae were gradually transferred to PBST through 10 min washes at room temperature in 75% 2 X SSC / 25% PBST, 50% 2 X SSC / 50% PBST, 25% 2 X SSC / 75% PBST, and 100% PBST. To reduce non-specific binding of antibody, larvae underwent preadsorption in block solution (2% calf serum, 2 mg/mL bovine serum albumin, PBST, filtered), and shaken at room temperature for 2 hours. For antibody adsorption, larvae were left overnight at 4 °C in 1/1000 diluted antibody (anti-digoxigenin-aP Fab fragments, Roche cat #15421-019, in block solution). To remove unbound antibody and block solution, larvae were given seven 15 min washes in PBST. They were subsequently given three 5 min washes in phosphatase alkaline buffer (PAB, 100 mM TrisHCl pH 9.5, 50 mM $MgCl_2$, 100 mM NaCl, 0.1% tween 20 and sterile H_2O). Larvae were placed in a solution of PAB, 226 µg/mL nitroblue terazolium (NBT), and 175 µg/mL 5-bromo-4-chloro-3-indolyl phosphate (BCIP). The alkaline phosphatase enzyme which is conjugated to the anti-DIG antibody reacts with these chemicals to produce an indigo precipitate localized within the larvae. This colouration reaction is light sensitive and is done with gentle agitation. To avoid over-colouration, larvae are initially checked every 20 min, and then every 10 minutes until the desired colouration was achieved. To stop the colouration reaction, larvae were washed 3 times for 5 min with 1 mM EDTA in PBST. They were then post-fixed in 4% PFA overnight at 4 °C to prevent degradation. Larvae were gradually equilibrated into 100% glycerol, which removes some background staining and facilitates positioning during imaging.

Table 1: List of primers used to amplify and subclone enhancer sequences from BACs. All primers also have 5' SalI restriction endonuclease sites.

Enhancer	Forward Primer	Reverse Primer
M I12b	5' GTCGACCGTACAGCTGCAAACCCAAGA	5' GTCGACAGAGGATATTAAAGAGGTATCT
M I56i	5' GTCGACTCAGTCTTGTCATTTTCTAGC	5' GTCGACCTGCAGCCTCTTCCATTCTT
SC I12b	5' GTCGACTTCTGCCAAAAGCTCCAAAT	5' GTCGACTTGCAATGGTTGACATCTCTG
SC I56i	5' GTCGACGCCATGGGTCTGATCTCATT	5' GTCGACTCAGCTTGGCACTTTCACTG

Table 2: List of morpholino oligonucleotides

MO Target gene	Translation blocking MO sequence
ascl1a	5’ AAGGAGTGAGTCAAAGCACTAAAGT 3’
dlx1a	5’ CTCGCTCTCGCTCTCTGTACTGGTA 3’
dlx2a	5’ CTCCAGTCATGTTTTTCATACCGCA 3’
dlx5a	5’ TCCTTCTGTCGAATACTCCAGTCAT 3’
dlx6a	5’ TGGTCATCATCAAATTTTCTGCTTT 3’

3. RESULTS

3.1 Comparative analysis of intergenic enhancer constructs

One approach to identifying how *Dlx* genes have contributed to lineage-specific evolutionary changes in developmental processes is to better characterize the activity of their regulatory sequences. Species-specific temporal and spatial differences in developmental gene expression can underlie the divergence of adult phenotypes, and can arise due to the altered activity of regulatory elements such as enhancers. By experimentally comparing the activity of orthologous enhancer sequences of two different species to drive a reporter gene in identical developmental contexts, it may be predicted whether the enhancers have undergone functional changes or instead have highly conserved regulatory functions.

Plasmids containing a beta-globin minimal promoter upstream of *eGFP* and either dogfish (sc) I12b, I56i, or mouse (m) I12b, I56i constructs were co-microinjected with Tol2 RNA into single-cell zebrafish embryos. The beta-globin minimal promoter is inactive unless the nearby enhancer is activated. The purpose of this was to produce transgenic lines each expressing *eGFP* in a pattern recapitulating the endogenous activity of the enhancer. Subsequently, comparisons were to be drawn between scI12b / mI12b and scI56i / mI56i enhancer activity in the zebrafish forebrain, with the previously established lines, *Tg(6kbdlx1a/2aIG:GFP)* and *Tg(1.1kbI56i:GFP)* (Macdonald *et al.,* 2010; Yu *et al*., 2011). Approximately 30-40% of microinjected embryos showed reporter expression beginning at 24 hpf (n >100), but due to fish death and other problems, only one transgenic line was produced: scI12b. Unfortunately, when scI12b F1 individuals were bred, reporter expression was lost in the

F2 offspring. Because of this, only observations on transient reporter gene expression could be made.

The zfI12b enhancer can drive reporter gene expression in forebrain regions that endogenously express *dlx1a/dlx2a,* including the telencephalon and the diencephalon (prethalamus) (**Figure 10)**. A similar proportion (30-40%, n >100) of embryos injected with scI12b and mI12b drove reporter gene expression in the same brain regions beginning around 24 hpf and continuing until 52 hpf. As eGFP expression in primary transgenic fish is highly mosaic, it is difficult to accurately compare the distinct boundaries of eGFP expression in the forebrain, the intensity / number of cells expressing eGFP, and the temporality of its expression. As the larvae develop, the injected plasmid becomes continuously more diluted and therefore there is a time limit for visibility in transient reporter expression. Nonetheless, it is possible to determine that the dogfish I12b enhancer drives reporter expression in similar brain regions and developmental time points as both the zfI12b enhancer and endogenous zebrafish *dlx1a/dlx2a* genes.

In the *Tg(1.1kbI56i:GFP)* line, the zebrafish I56i enhancer appears to recapitulate endogenous *dlx5a/dlx6a* expression in the forebrain (telencephalon, diencephalon), but not in the midbrain (optic tectum) (Yu *et al.*, 2011). The timing of reporter expression also coincides with endogenous gene expression in this line. Larvae injected with the dogfish I56i and mouse I56i constructs also showed reporter expression specifically in the forebrain at time points between 24-48 hpf (30-40%, n > 100 for each category) (**Figure 11)**. For both I12b and I56i constructs, sequence orientation had no visible effect on reporter expression.

Figure 10 : Dogfish and mouse I12b enhancers can drive reporter expression in similar zebrafish forebrain regions as endogenous zebrafish I12b. *Dlx2a* is expressed in the telencephalon and diencephalon of the zebrafish forebrain at 24 hpf (A) and 48 hpf (B) (in situ hybridization). The transgenic line with a 6 kb *dlx1a/dlx2a* intergenic region including the I12b enhancer, shows eGFP expression (C and D) that closely mimics endogenous *dlx* expression pattern. In zebrafish embryos injected with the dogfish 502 bp I12b plasmid (E - H), transient reporter expression can be seen in forebrain regions endogenously expressing *dlx1a/dlx2a.* Similarly, in zebrafish embryos injected with the mouse 426 bp I12b plasmid (I - L), transient reporter expression can be seen in forebrain regions that endogenously express *dlx1a/dlx2a.* F, H, J, and L are frontal views, all others are lateral. Scale bar, 25 μm N >100.

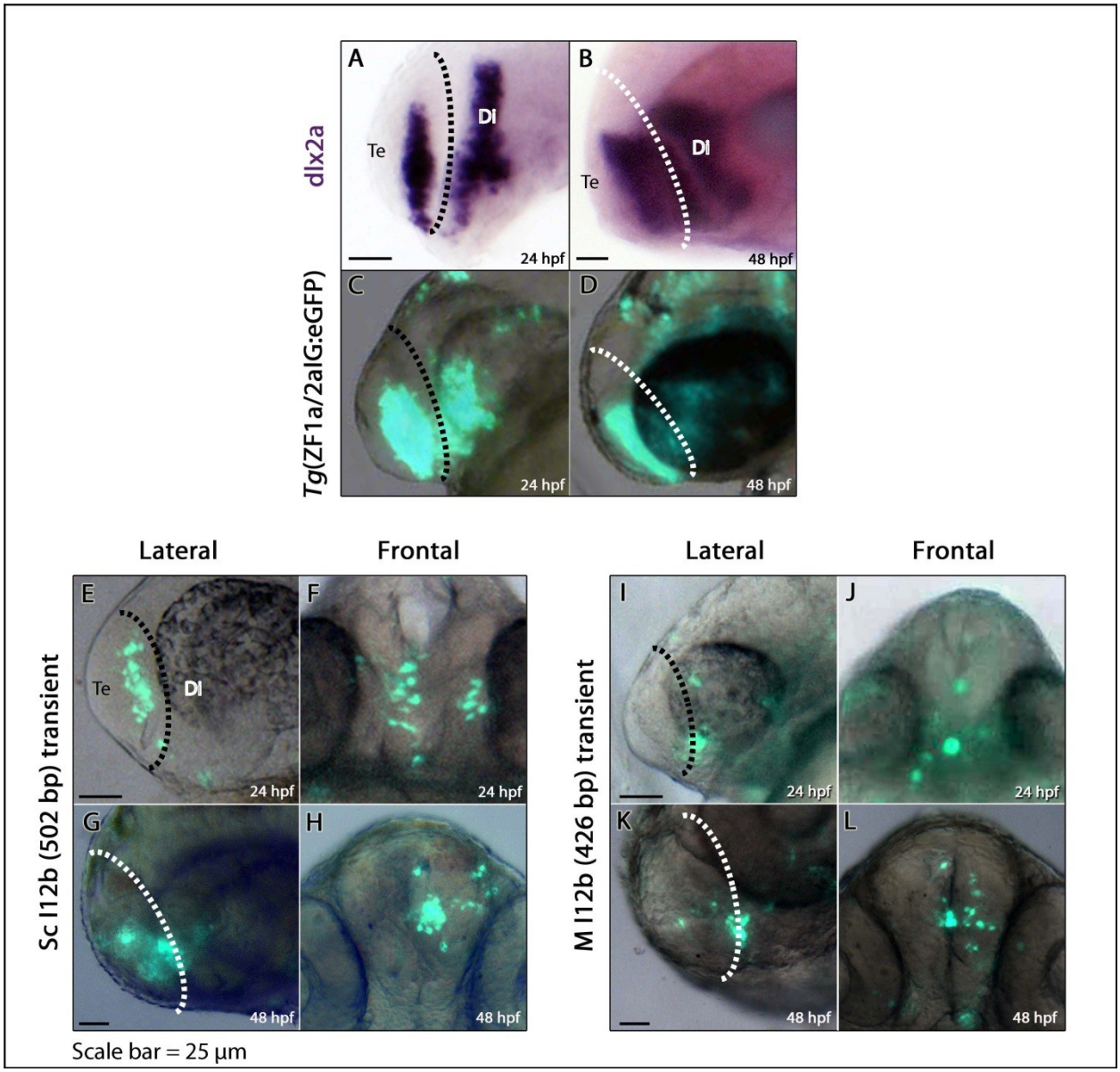
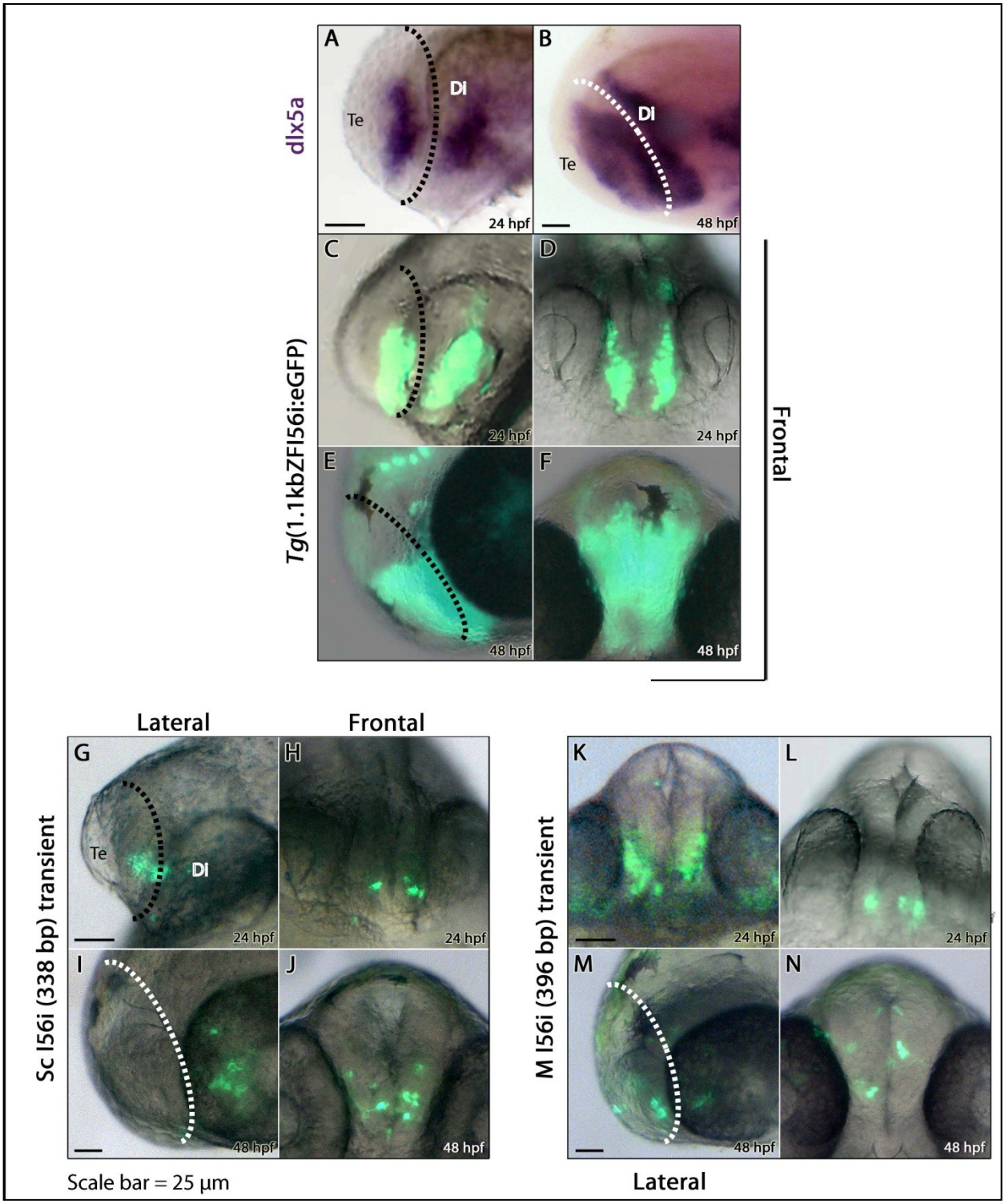

dlx2a
Tg(ZF1a/2aIG:eGFP)
A
B
C
D
Te
Dl
24 hpf
48 hpf
Lateral
Frontal
E
F
G
H
I
J
K
L
Sc I12b (502 bp) transient
M I12b (426 bp) transient
Scale bar = 25 µm

Figure 11: Dogfish and mouse I56i enhancers can drive reporter expression in similar zebrafish forebrain regions as endogenous zebrafish I56i. *Dlx5a* is expressed in the telencephalon and diencephalon of the zebrafish forebrain at 24 hpf (A) and 48 hpf (B) (in situ hybridization). The transgenic line with a 1.1 kb *dlx5a/dlx6a* intergenic region including the I56i enhancer, shows eGFP expression (C - F) that closely mimics endogenous *dlx* expression pattern. In zebrafish embryos injected with the dogfish 338 bp I56i plasmid (G - J), transient reporter expression can be seen in forebrain regions endogenously expressing *dlx5a/dlx6a.* Similarly, in zebrafish embryos injected with the mouse 396 bp I56i plasmid (K - N), transient reporter expression can also be seen in forebrain regions that endogenously express *dlx5a/dlx6a.* Panels K and L show two embryos representing the range in intensity of eGFP expression (similar range found in with dogfish constructs. D, F, H, J, K, L and N are frontal views, all others are lateral. Scale bar 25 μm. N > 100


dlx5a
A
B
Dl
Te
24 hpf
48 hpf
Tg(1.1kbZFI56i:eGFP)
C
D
E
F
Frontal
Lateral
Frontal
Sc I56i (338 bp) transient
G
H
I
J
M I56i (396 bp) transient
K
L
M
N
Scale bar = 25 µm
Lateral

3.2 Morpholino oligonucleotide targeted down-regulation of *ascl1a* and *dlx* genes

The gene regulatory network involving *Mash, Dlx1/Dlx2, Dlx5/Dlx6* and *Gad* plays a central role in the differentiation and tangential migration of GABAergic interneurons in the mouse forebrain, but has not been well characterized in the zebrafish. Taking into consideration the inherent differences in anterior neural tube folding between zebrafish and mice (evagination versus eversion), and the resulting differences in forebrain structure, the orthologous genes in zebrafish seem to be expressed in analogous domains (Macdonald *et al.,* 2010). However, to illuminate whether the coinciding expression domains reflect functional and regulatory relationships between these genes, morpholino oligonucleotides were used to target down-regulation of the *ascl1a, dlx1a/dlx2a,* and *dlx5a/dlx6a* genes. Subsequently, *in situ* hybridization was used to qualitatively examine expression patterns of genes downstream to those down-regulated. In mice, the MASH protein is able to bind the *Dlx1/Dlx2* intergenic I12b enhancer (Poitras *et al*., 2007). Furthermore, *Mash* $^{-/-}$ targeted mutants mis-express *Dlx* genes in the ganglionic eminences and exhibit a decrease in *Gad67* expression in the same region. To verify that the microinjected material was successfully taken into embryonic cells, 'sham' eGFP protein was injected into one-cell embryos, and larvae were visualized with UV light at 24 hpf and found to have uniform eGFP fluorescence.

A morpholino oligonucleotide (MO) that blocks the translation of ascl1a transcripts was microinjected into one-cell embryos. These embryos were fixed at 24 hpf and 48 hpf and subjected to *in situ* hybridization using complementary RNA probes for *dlx1a, dlx2a, dlx5a* and *gad1a*. In *ascl1a* morphants, telencephalic and dorsal diencephalic expression of the *dlx* genes was not visibly affected. However, in the prethalamus and possibly hypothalamus, there is a visible decrease in *dlx* expression. This can be seen particularly at 48 hpf. Prethalamic

downregulation of *dlx2a* in *ascl1a* morphants is shown in **Figure 12.** Results were identical between *dlx1a* and *dlx2a,* so only those of *dlx2a* are shown. Prethalamic downregulation of *dlx5a* in *ascl1a* morphants is shown in **Figure 13.** Results were not gathered for *dlx6a* expression, as a probe sufficiently specific for this transcript has not been produced. *Gad1a* expression is relatively unaffected in the telencephalon and dorsal diencephalon of *ascl1a* morphants, but similarly shows a marked decrease in prethalamic expression. As both genes in a *dlx* bigene pair are thought to have highly redundant functions, it was necessary to down-regulate each one to accurately characterize possible downstream effects. MOs targeting *dlx1a* and *dlx2a* were co-injected into embryos, which were fixed at 24 hpf and 48 hpf, and subjected to *in situ* hybridization using *dlx5a, dlx6a* and *gad1a* probes. The most effective way to verify that a MO has successfully decreased the translation of its target transcript is to use immunohistochemistry to visualize a decrease in protein levels. Unfortunately, effective antibodies binding specifically to zebrafish dlx proteins have not yet been discovered. One method to determine if *dlx* genes have been down -regulated takes advantage of the pleiotropic activity of *dlx1a* and *dlx2a.* These genes play a role in the proper development of zebrafish branchial arches. In 5 dpf *dlx1a/dlx2a* double morphants, there are notable defects in the growth and patterning of the cartilaginous precursors of branchial arch structures (Sperber *et al*., 2008). To ensure that the *dlx1a* and *dlx2a* MOs were properly down-regulating their target proteins, the branchial arch phenotype was replicated (**Figure 14)**. Approximately 70% (46/65) of 5 dpf *dlx1a/dlx2a* double morphants showed moderate to severe defects in branchial arch development, whereas only 3% (2/59) of larvae injected with control MO had this phenotype. Because there is a lethal limit to the quantity of MOs taken into an embryo, it becomes diluted as cells multiply during development. By the stage of 5 dpf, the effect of MOs is negligible as the number of a given cell's endogenously

Figure 12: *Dlx2a* expression is reduced in the ventral thalamus in *ascl1a* zebrafish morphants. In situ hybridization using a probe targeting *dlx2a* transcripts of control morpholino oligonucleotide - injected (panels A, B and C), and *ascl1a* translation blocking MO - injected larvae (panels D, E and F). Asterisk indicates down-regulated *dlx2a* in the ventral prethalamus and possibly hypothalamus in 24 hpf (panel D) and 48 hpf (panels E and F) (63%, n = 72/114). Arrowheads indicate visibly unaltered telencephalic expression of *dlx2a* between the two treatments. Scale bar 25 μm.

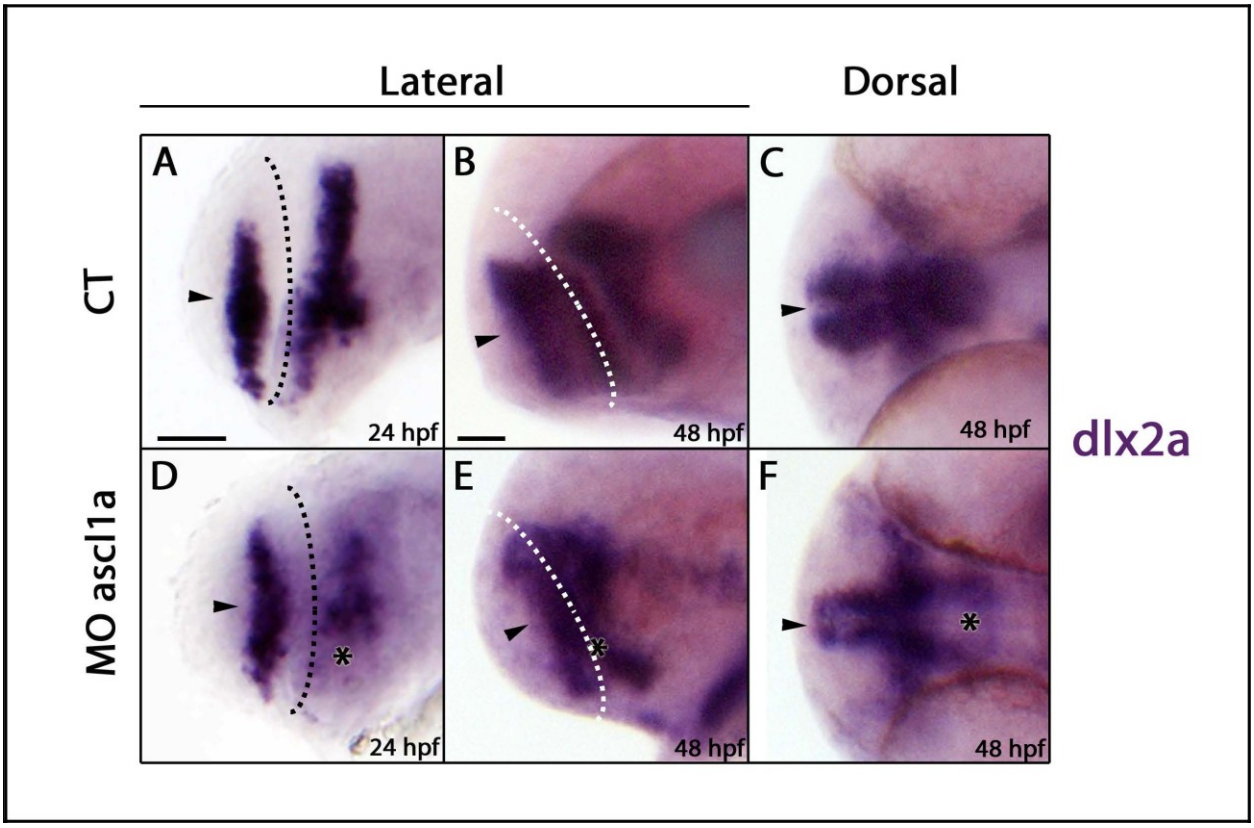

Lateral
Dorsal
CT
MO ascl1a
A
B
C
D
E
F
24 hpf
48 hpf
48 hpf
24 hpf
48 hpf
48 hpf
dlx2a

Figure 13: *Dlx5a* expression is reduced in the ventral thalamus in *ascl1a* zebrafish morphants. In situ hybridization using a probe targeting *dlx5a* transcripts of control morpholino oligonucleotide - injected (panels A, B and C), and *ascl1a* translation blocking MO - injected larvae (panels D, E and F). Asterisk indicates down-regulated *dlx5a* in the ventral prethalamus and possibly hypothalamus in 24 hpf (panel D) and 48 hpf (panels E and F) (56%, n = 120). Arrowheads indicate relatively unaltered telencephalic expression of *dlx5a* between the two treatments. To show that the latter expression pattern does not reflect an earlier stage of larval development (caused by developmental delay); pictures of 36 hpf larvae are included (panels G and H). Scale bar is 25 μm.

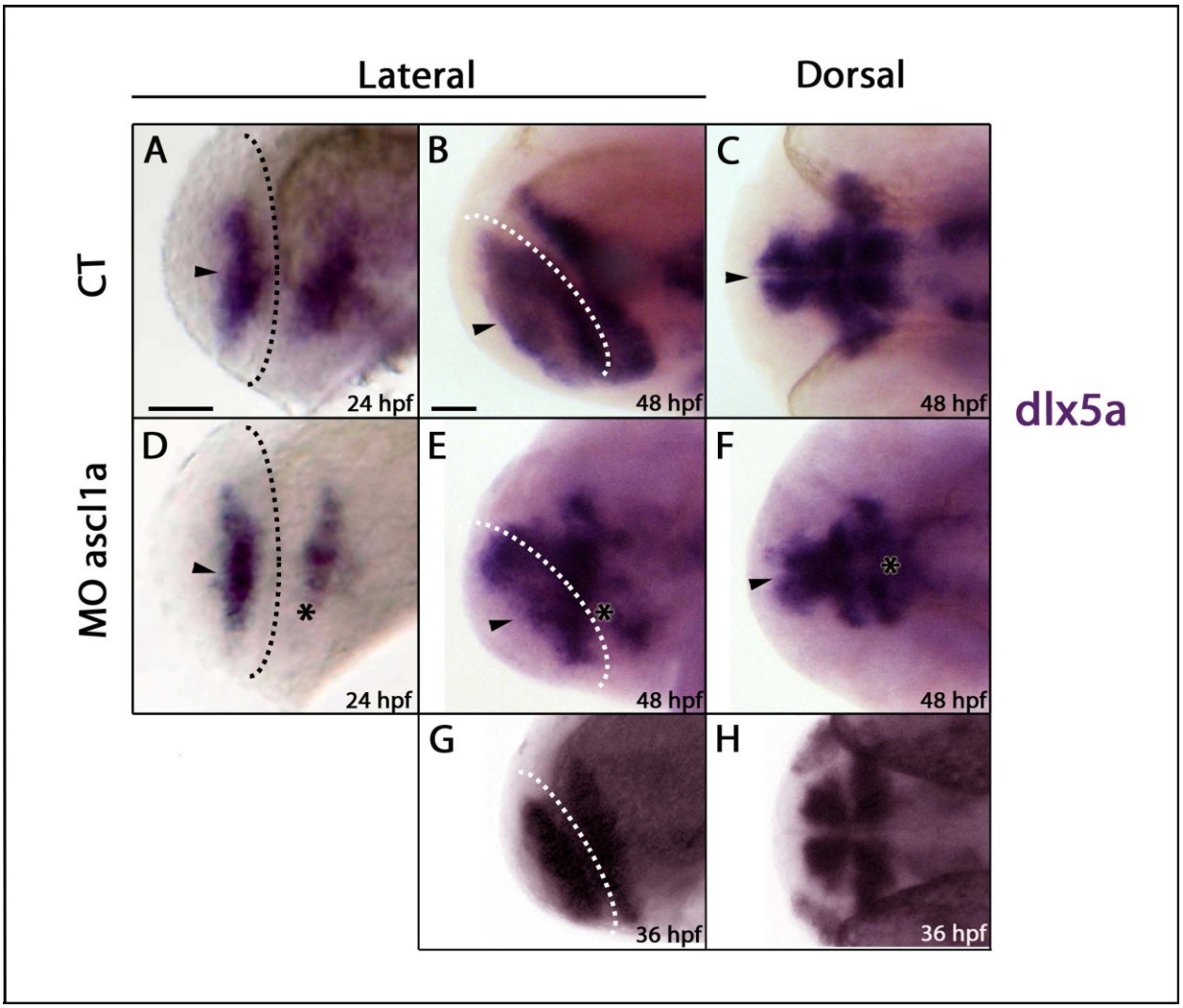
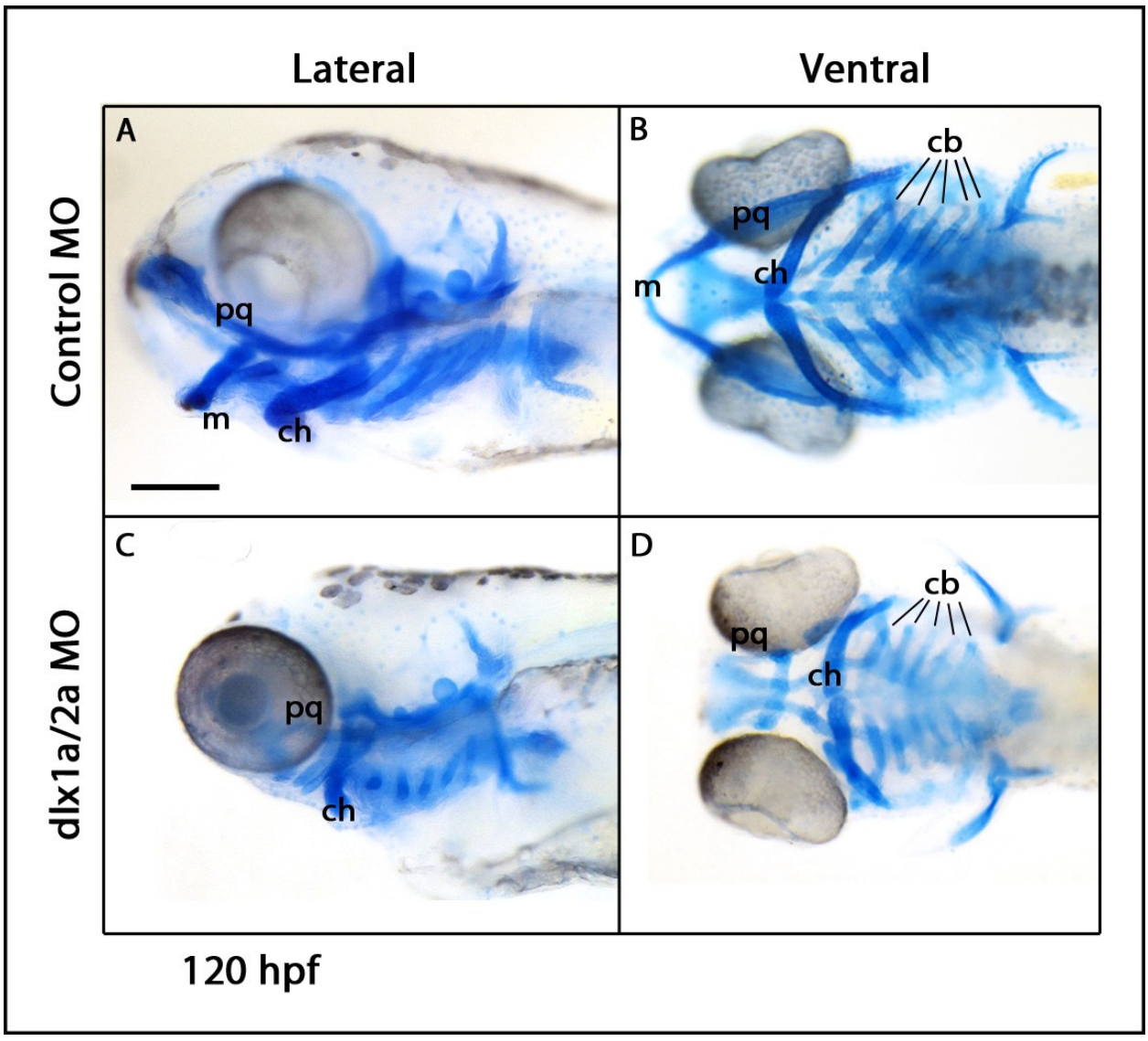

Lateral
Dorsal
CT
MO ascl1a
A
B
C
D
E
F
G
H
24 hpf
48 hpf
48 hpf
24 hpf
48 hpf
48 hpf
36 hpf
36 hpf
dlx5a

Figure 14: Mis-patterning of zebrafish craniofacial structures in *dlx1a/dlx2a* morphants. Cartilaginous structures are stained with alcian blue. In control morpholino oligonucleotide - injected embryos (panels A and B), the craniofacial structures of 5 dpf control –MO injected zebrafish are normal (5 % malformation, n = 59). In *dlx1a/dlx2a* (panels C and D) double morphants, however, there is malformation of the Meckel's cartilage (m), palatoquadrate (pq) and ceratohyal (ch), but not in the ceratobranchials (cb) (70% malformation, n = 93/(65)).
Scale bar 100 μm.


Lateral
Ventral
Control MO
dlx1a/2a MO
A
B
C
D
pq
m
ch
cb
120 hpf

produced target transcripts vastly outnumbers the MO molecules in the vicinity. It is assumed here, therefore, that a 70% rate of mutated branchial arch phenotype in 5 dpf *dlx1a* and *dlx2a* morphants signifies an even higher MO efficiency at earlier stages (24 hpf and 48 hpf). Lacking antibodies against Dlx1a and Dlx2a, this information indirectly shows that the MOs have effectively decreased the translation of these proteins.

In these morphants, there is not a detectable change in *dlx5a* expression in the forebrain. Similarly to *ascl1a* morphants, *gad1a* expression does not change noticeably in the telencephalon, but is down-regulated in the prethalamus at 48 hpf. In mice, *Dlx5* and *Dlx6* are partially activated by *Dlx1* and *Dlx2*, and contribute to the correct migration and differentiation of Gad-expressing neurons. To determine if dlx5a and dlx6a proteins play an equivalent role in activating *gad1a* in zebrafish, embryos were microinjected with *dlx5a* and *dlx6a* MOs and *gad1a* expression was examined at 24 hpf and 48 hpf. There were no detectable differences in *gad1a* expression in these morphants at either stage **(Figure 15** and **Figure 16)**.

These data shows that (1) the down-regulation of Ascl1a activity leads to decreased *dlx1a, dlx2a, dlx5a* and *gad1a* expression in the ventral diencephalon, and 2) down-regulation of *dlx1a* and *dlx2a* activity does not noticeably change *dlx5a* expression in the forebrain, but decreases *gad1a* expression in the ventral diencephalon.

3.3 Exogenous expression of *dlx* genes in *ascl1a* and *dlx* morphants

The next thing to investigate was whether *ascl1a* and *dlx1a/dlx2a* act in the same or parallel cascades in regulating the expression of *gad1a* in this region. If exogenous expression of *dlx1a* or *dlx2a* could rescue the *gad1a* phenotype in *ascl1a* morphants, this would signify that these genes act in the same genetic cascade. Conversely, if the phenotype is not rescued this

Figure 15: The expression of *gad1* is lost in the prethalamus but not the telencephalon in *ascl1a* single and *dlx1a/dlx2a* double morphants at 48hpf. *In situ* hybridization on whole mount embryos. (A) Expression of *gad1a* in control morpholino injected embryos is present in the telencephalon (Te), prethalamic (Pt) and hypothalamic (Hy) diencephalon (dorsal view in E). (B) Morpholino knockdown of *ascl1a* results in a loss of prethalamic (dashed box) and hypothalamic (asterisk) *gad1* expression (dorsal view in F). (C) Double morpholino knockdown of *dlx1a* and *dlx2a* results in decreased prethalamic and hypothalamic *gad1* expression (dorsal view in G). (D) In *dlx5a/dlx6a* there is a slight reduction of prethalamic and hypothalamic *gad1a* expression (dorsal H). A-D are lateral views, dorsal is up, E - H are dorsal views. Anterior is to the left. Dashed line indicates telencephalon-diencephalon boundary. Scale bar, 50μm. N > 100 per treatment.

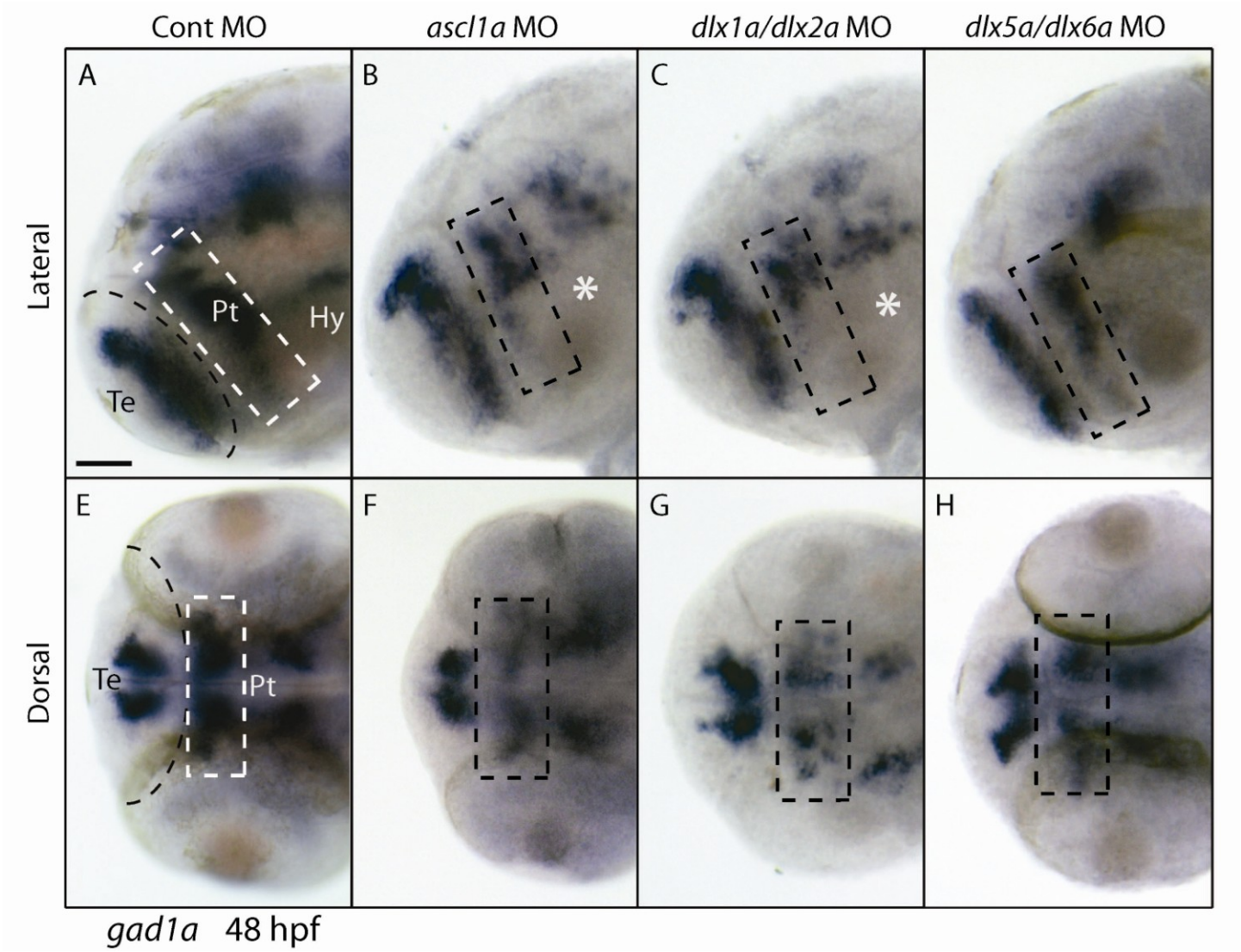
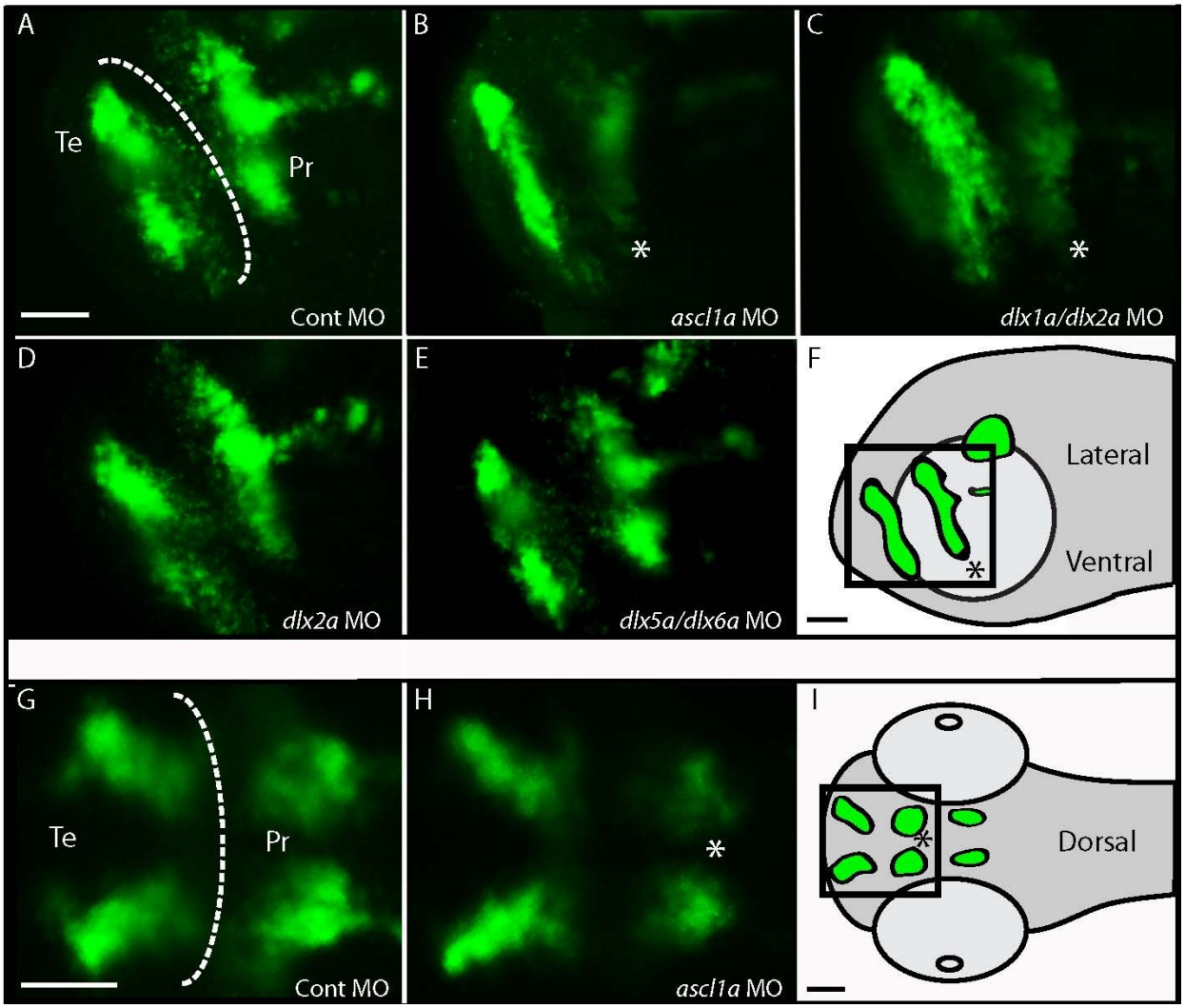
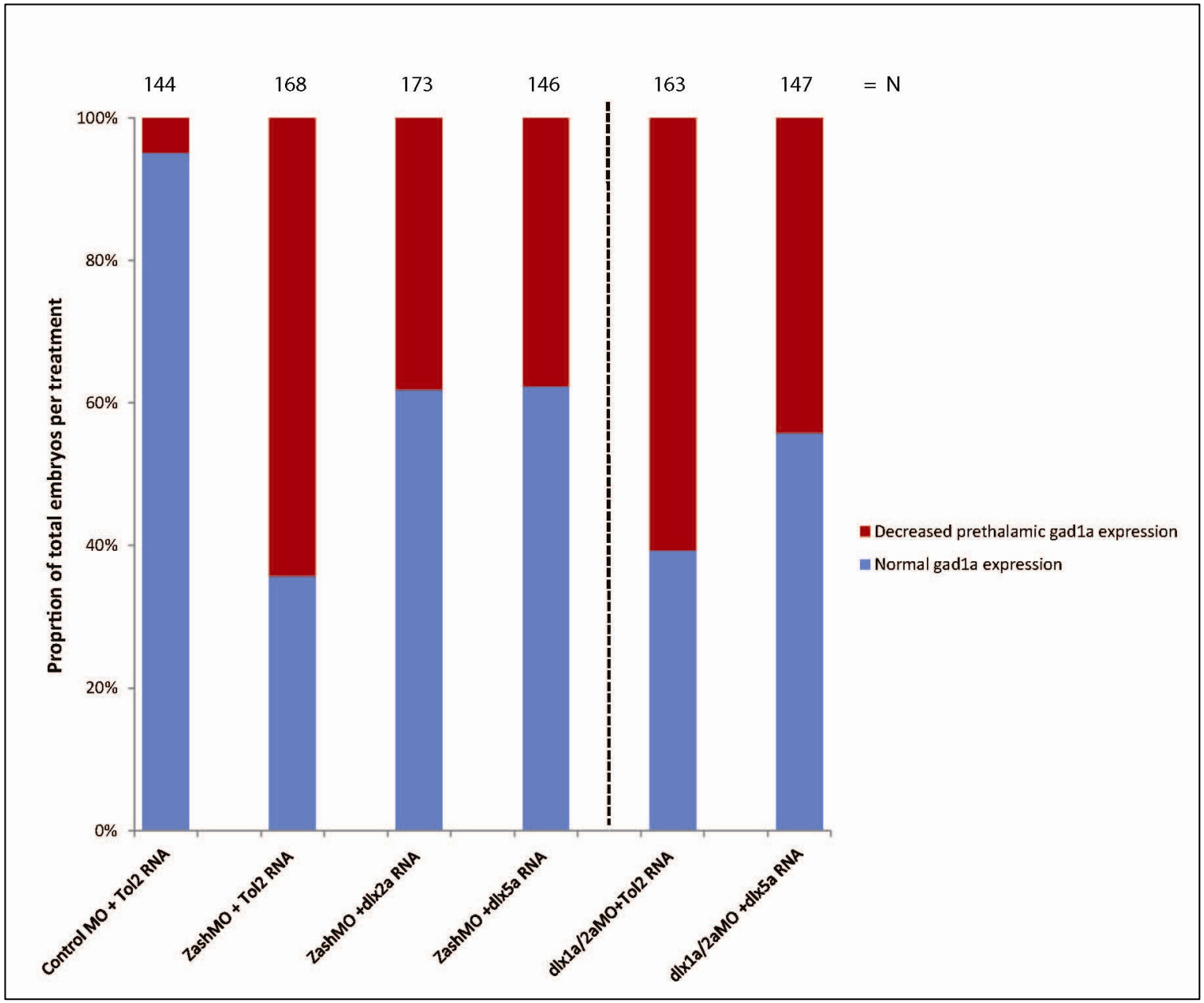

Cont MO
ascl1a MO
dlx1a/dlx2a MO
dlx5a/dlx6a MO
Lateral
Dorsal
A
B
C
E
F
G
H
Pt
Hy
Te
Te
Pt
*
*
gad1a 48 hpf

Figure 16: Diencephalic expression of *gad1a* is reduced in 48 hpf *ascl1a* single and *dlx1a/dlx2a* double morphants. Confocal images of fluorescent *in situ* hybridization in whole mount embryos. (A) Control morpholino injected embryos have normal telencephalic and diencephalic expression of *gad1*. (B) *Ascl1a* morphants and (C) *dlx1a/dlx2a* double morphants have decreased prethalamic *gad1* expression (asterisk). Single *dlx2a* (D), and *dlx1a* (data not shown), morphants do not have this *gad1* phenotype. Similarly, single *dlx5a* and *dlx6a* morphants (data not shown), and *dlx5a/dlx6a* double morphants (E) have normal *gad1* expression. (F) A diagram indicating lateral field of view of confocal images A-E. Dorsal view of forebrain in (G) control morpholino injected fish and (H) *ascl1a* morphants (asterisk indicates decreased prethalamic *gad1* expression). Anterior is to the left. Scale bar = 25μm.


A
Te
Pr
Cont MO
B
*
ascl1a MO
C
*
dlx1a/dlx2a MO
D
dlx2a MO
E
dlx5a/dlx6a MO
F
Lateral
Ventral
*
G
Te
Pr
Cont MO
H
*
ascl1a MO
I
Dorsal
*

would signify that *ascl1a* and *dlx1a/dlx2a* operate in distinct but parallel cascades regulating *gad1a* expression. Rescue experiments were carried out by co-injecting *ascl1a*-MO with *dlx2a* or *dlx5a* mRNA, and performing *in situ* hybridization for *gad1a* in 48 hpf larvae. As both genes in a *Dlx* bigene pair have highly redundant expression patterns and biochemical function, it was deemed adequate to inject mRNA from only one gene of each pair. The *gad1a* phenotype of a given individual was classified as being normal (resembling the wild type expression), or down-regulated in the ventral diencephalon (**Figure 15,** see panels A and E for normal *gad1a* expression, and panels B and F, for example of down-regulation in diencephalon**)**. A cohort of larvae for a given treatment would have varying combinations of these two phenotypes. Exogenous expression of *dlx2a* and *dlx5a* mRNA partially and similarly rescued the down-regulated *gad1a* phenotype seen in *ascl1a* morphants (**Figure 17)**. To a lesser extent, exogenous *dlx5a* expression rescued the *gad1a* phenotype seen in *dlx1a/dlx2a* double morphants. *Gad1a* phenotypic categorization was agreed upon by several lab members, and scoring of treatments was done double blind.

Figure 17: Partial rescue of *gad1a* diencephalic down – regulation by *dlx2a* and *dlx5a* exogenous expression in *ascl1a* and *dlx1a/dlx2a* morphants. X –axis is treatment of material injected into embryos. Y –axis is proportion of total embryos per injected treatment with normal (blue), or down-regulated (red), diencephalic *gad1a* phenotype. Between the control MO + Tol2 injected and *ascl1a*MO + Tol2 RNA injected treatments, there is a dramatic change in normal phenotypes (95%, n = 144, to 36%, n = 168, respectively). The latter is thereafter treated as the baseline for comparison with *ascl1a*MO + Dlx2a RNA as well as *ascl1a*MO + Dlx2a RNA treatments. Both the exogenous expression of Dlx2a and Dlx5a increase the proportion of embryos per treatment with the normal *gad1a* diencephalic phenotype (62%, n = 173, and 63%, n = 146, respectively). Similarly, Dlx5a exogenous expression (56% normal, n = 163), can partially rescue the *gad1a* phenotype in *dlx1a/dlx2a* double morphants (39% normal, n = 147).


144
168
173
146
163
147
= N
100%
80%
60%
40%
20%
0%
Proprtion of total embryos per treatment
Decreased prethalamic gad1a expression
Normal gad1a expression
Control MO + Tol2 RNA
ZashMO + Tol2 RNA
ZashMO +dlx2a RNA
ZashMO +dlx5a RNA
dlx1a/2aMO+Tol2 RNA
dlx1a/2aMO +dlx5a RNA

DISCUSSION AND PERSPECTIVES

4.1.1 Multiple layers of developmental gene regulation are susceptible to mutational variation

Adult organismal form is determined by spatiotemporally and hierarchically organized gene regulatory networks that drive morphogenesis. Morphological evolution arises from the fixation in a population of genetically inherited developmental novelties that bestow new fitness-changing phenotypes upon an organism. Even though for simplicity I am here ignoring the importance of epigenetic, post-transcriptional, post-translational, and chromatin-level genome regulation, there is still a striking range of mechanisms by which these developmental novelties can be introduced, at every level of genetic and developmental complexity.

The fundamental building blocks of GRNs are signalling and transcription factor proteins, the *cis*-regulatory element DNA sequences they bind to, and the genes that are thereafter transcriptionally controlled. The *cre*-mediated combinatorial code of transcription regulation, in which combinatorial binding of transcription factor complexes can alter the regulatory capacity of enhancers, adds another layer of complexity. Similarly, a transcription factor encoded by a given *cis*-regulated gene may serve a variety of different purposes (i.e. activation or repression), dependent on the conditional presence or absence of co-factors. Due to the progressive nature of organismal development, GRNs involved during morphogenetic events are inherently transient and may only exist at specific spatiotemporal zones. Input changes to these GRNs can alter the spatiotemporal zones in which they are activated. A few transcription factor inputs can initiate the activation of an entire GRN that underlies the patterning of a given limb or structure. Thus, changes in the regulation of one gene may have much larger

consequences. Genetic mutations can bring about changes at each of these levels of developmental regulation: (1) Mutations in, or losses of, *cis*-regulatory elements can lead to different binding affinities with transcription factors; (2) non-synonymous mutations in transcription factors may induce changes in biochemical properties, leading to changes in the binding affinity with certain DNA sequences or co-factors; (3) when the gene regulatory changes produced by (1) and (2) occur for the crucial initial inputs of larger downstream GRNs, entire developmental processes can be affected. Furthermore, mutation-induced changes in gene regulation may cause the architecture of GRNs to change. Regulatory changes of a node that was previously important in a GRN (with high connectivity), may cause it to be lost or replaced in the network. Conversely, regulatory changes of a node that was previously not involved in a GRN may result in its co-option into the network. The main point here is that there are many ways through which genetic mutations may affect gene regulation and organismal development.

Unfortunately, the simple drawing of comparisons between fully sequenced genomes is inadequate for illuminating how species have morphologically diverged. Instead, experimental approaches can take advantage of the modular nature of nodes and GRNs, and artificially 'tinker' with the very GRNs that control the development of diverged morphological structures. In this way, we can get closer to determining how developmental processes, GRNs, and individual genes were likely to have undergone lineage-specific changes in the past resulting in divergent phenotypes. It is to these ends that zebrafish are used for their large egg size, sequenced genome, transparent embryonic development, and high amenability to genetic manipulation.

In the present work, the first set of experiments was designed to make cross-species comparisons of the regulatory activity of orthologous enhancers of *Dlx* genes. This examined the most basic level of developmental regulation, as (1) described above. The second set of

experiments was designed to better characterize the regulatory relationships between several developmental genes in the zebrafish, the orthologs of which are known to contribute to mouse forebrain GABAergic interneuron development. Unlike the first experiments, this was a GRN-level approach to investigating and comparing developmental gene regulation between species.

4.1.2 Mouse and dogfish I12b and I56i enhancers drive reporter expression in the zebrafish forebrain

In vertebrates, the *Dlx* genes encode a family of homeodomain transcription factors implicated in a variety of developmental processes. In the early mouse and zebrafish forebrain *Dlx1/Dlx2* (*dlx1a/dlx2a*) and *Dlx5/Dlx6* (*dlx5a/dlx6*) bigene clusters are expressed in the telencephalic ventricular zone and in the diencephalic thalamus. In dogfish, *Dlx5/Dlx6* are expressed slightly before *Dlx1/Dlx2* in the telencephalon, and all four genes are expressed in some cells in the diencephalon (personal communication from Melanie Debiais-Thibaud). The ability for *Dlx* genes to be expressed in very different tissues during development is facilitated by the presence of multiple CREs near or within these bigene clusters. I12b is a conserved non-coding element located in the *Dlx1/Dlx2* intergenic region and acts as an enhancer for the transcription of its neighbouring genes in the forebrain. Similarly, the conserved I56i sequence is located in the *Dlx5/Dlx6* intergenic region and is also considered to drive the forebrain expression of its neighbouring genes. Zebrafish I12b- and I56i- containing sequences also drive reporter expression in zebrafish forebrain regions that endogenously express *dlx* genes. Between the three species, homologous enhancers share around 80-85% sequence similarity, but it is not known whether the 15-20% sequence divergence coincides with differences in how these enhancers bind transcription factors. In this work I determined that both the dogfish and mouse

I12b and I56i enhancers are able to, independent of orientation, drive reporter expression specifically in the zebrafish forebrain. Also, the dogfish I12a and I56ii sequences did not activate reporter expression in the zebrafish forebrain. This signifies that in a homogeneous genetic background (i.e. the zebrafish), the dogfish and mouse enhancers show similar regulatory activity in the forebrain to those of zebrafish. Previous work has shown that in their respective species, mouse and zebrafish enhancers drive reporter gene expression in analogous brain regions, and are likely bound by similar transcription factors (Poitras *et al.*, 2007; Yu *et al.*, 2010). However, there are two reasons why we cannot conclude from these data that the functions of these enhancers are also conserved in dogfish: (1) There is no information on which transcription factors bind endogenously to the dogfish enhancers (what are the nodal inputs?), and (2), we do not know *if* or *where* (or *when*) the dogfish enhancers endogenously activate *Dlx* gene expression (what are the nodal outputs?). Although the enhancer sequences are highly conserved in dogfish, they could have hypothetically acquired novel transcription factor inputs, or different secondary co-factors, that bring about alternative regulatory properties. In other words, because enhancers work in concert with other factors in order to function, we cannot deduce from enhancer sequence conservation alone that function has also been conserved.

There are, however, several lines of reasoning suggesting that dogfish enhancers do in fact have a conserved function to those of their vertebrate cousins. Firstly, the highly overlapping expression patterns of both genes in a *Dlx* bigene cluster has been attributed to the simultaneous activation of both promoters by an enhancer-transcription factor complex forming in the intergenic region. In dogfish, *Dlx* gene pairs have overlapping expression patterns similar to those found in zebrafish and mice. This suggests that if one function of these enhancers in the latter two species is to coordinate similar expression of both neighbouring genes, then this

function is likely conserved in dogfish. Secondly, there is some variation among these species in where/when *Dlx* genes are expressed during development, however, *Dlx1/Dlx2* and *Dlx5/Dlx6* are all expressed in the anterior forebrain. As the I12b and I56i enhancers have been primarily found to regulate *Dlx* expression in this region (in mice and zebrafish), it would be a parsimonious assumption that this role has been conserved in dogfish. The lack of regulatory activity of dogfish I12a and I56ii in the forebrain is consistent with mice and zebrafish enhancers (in homogeneous and heterogeneous genetic backgrounds), and provides indirect evidence that the functions of I12b and I56i have also been conserved in dogfish. Thirdly, gene or genome duplication events give rise to the functional divergence of *cis*-regulatory elements (Teichmann and Babu, 2004). The multiplication of developmental transcription gene families, for example, provides the material for the subfunctionalization (partitioning of one ancestral gene function between two genes) and neofunctionalization (gain of new functions) within the gene family. This is one of the pivotal molecular phenomenons underlying the generation of diversity during metazoan evolution. After a gene or genome duplication event, *cis*-regulatory elements may experience relaxed selection, acquire mutations, and possibly begin responding to different regulatory states, thus altering their regulatory output. The lineage leading to teleosts is thought to have experienced an additional genome duplication event. We could therefore speculate that, all else being equal between mice, dogfish and zebrafish (teleost), zebrafish enhancers (in general) would have been the most likely to functionally diverge. But, since we already know that mouse and zebrafish I12b and I56i enhancers operate in highly conserved ways, it is predicted the dogfish enhancers will also have conserved regulatory functions.

4.1.3 Future directions for comparative functional analysis of *Dlx* enhancers in vertebrates

The experiments required to further this investigation include (1) the production of zebrafish transgenic lines containing dogfish enhancers upstream of a reporter gene (attempted in this work). There are limitations with analyzing transiently transgenic fish, as mosaicism prohibits accurate spatial and temporal characterization of reporter expression. It is possible that due to mosaicism, subtle inter-species differences in the intensity or spatiotemporal expression of eGFP were not observed. The activity of orthologous enhancers could more accurately be compared using fully transgenic zebrafish. (2) The activity of I12b and I56i enhancers must be studied endogenously in dogfish. This would allow us to determine whether their role as *Dlx* regulators in the forebrain has been conserved. Unfortunately, the dogfish has not yet been optimized for transgenesis (the embryo develops internally and is much less accessible for microinjection). Chromatin immunoprecipitation (ChIP) is a technique that could be used to determine which (if any), transcription factors bind endogenously to the dogfish I12b and I56i. MASH and ascl1a in mice and zebrafish, respectively, are among the transcription factors binding I12b, and Dlx is thought to bind I56i in mice and zebrafish.

4.2.1 Overview of early vertebrate diencephalon patterning GRNs

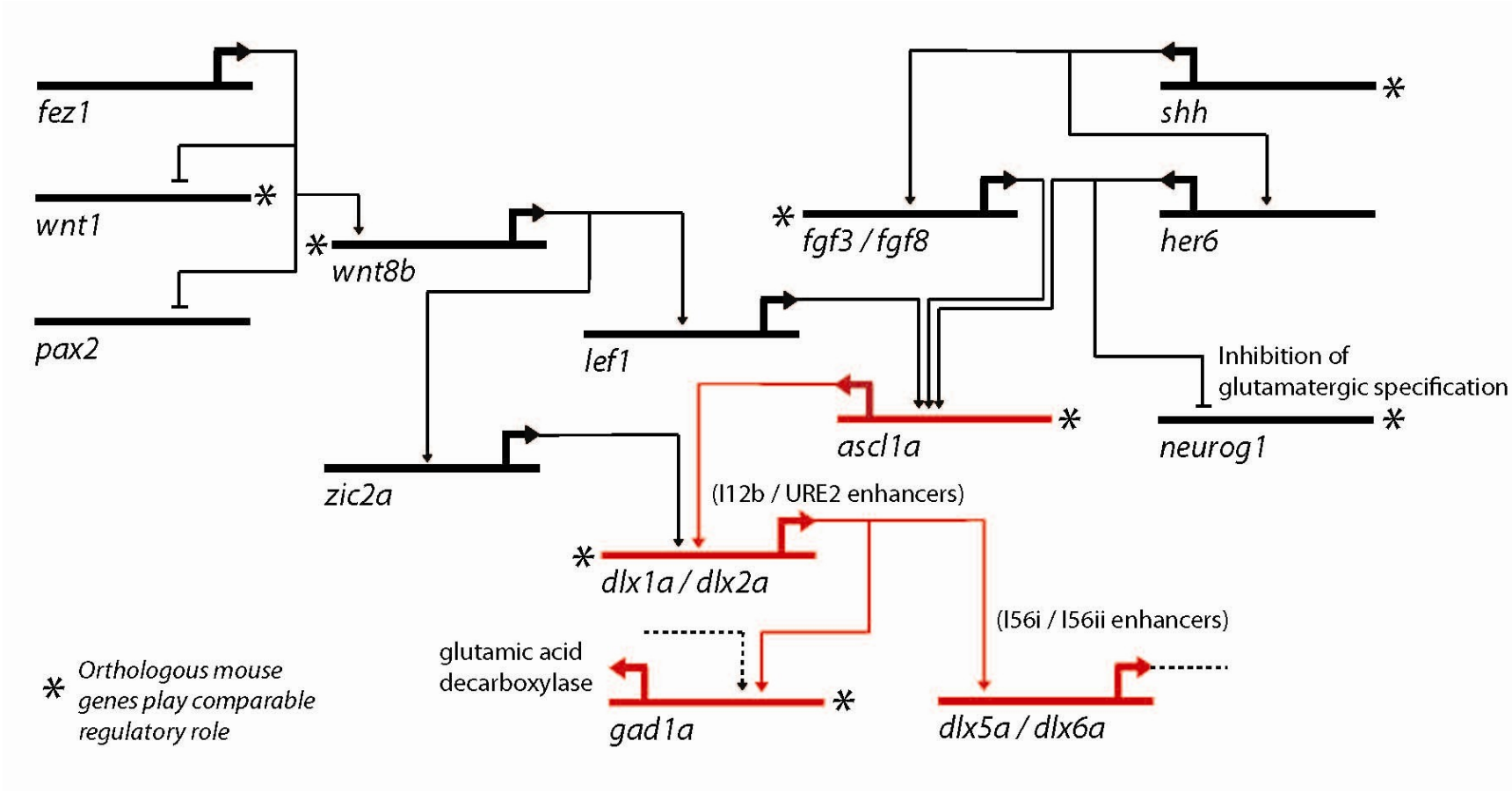
Dlx genes play a role in the timing of GABAergic interneuron precursor migration and differentiation in the ventral mouse forebrain. Here, the bHLH proneural transcription factor MASH is a regulator of *Dlx1/Dlx2* neural expression, which in turn plays a regulatory role in *Dlx5/Dlx6* expression, and ultimately leads to the activation of *Gad* genes (markers for cells with GABAergic interneuron fate).

In this work, I have shown that *gad1a* expression in the zebrafish developing prethalamus is dependent on *ascl1a* and *dlx1a/dlx2a* expression, whereas early telencephalic *gad1a* activity is independent of these genes. Furthermore, *dlx1a/dlx2a* expression is downstream of *ascl1a* in this region. To allow a better understanding of how these data fit in with the larger developmental context, the following paragraphs will outline the major genetic events that pattern the developing diencephalon.

The patterning of the diencephalon is determined by the mid-diencephalic organizer (MDO), which is also known as the zona limitans intrathalamica (ZLI). The correct formation and positioning of the MDO is dependent on early *sonic hedgehog* (*shh*) signaling and the expression of factors at the boundary between the prechordal (anteriormost) and epichordal neural plate regions (Scholpp and Lumsden, 2010). This boundary is located in the middle of the diencephalon and induces a dorsal extension to the primarily ventral, anteroposterior expression pattern of *shh* (Vieira *et al*., 2005; Guinazu *et al,* 2007). Fez transcription factors are expressed in the prechordal neural plate, with abutting expression to Otx transcription factors in the epichordal neural plate. Possibly through mutual cross-repression, these factors demarcate a sharp boundary in which *shh* expression is induced. In mouse and zebrafish, both Fez and Otx transcription factors are necessary for proper MDO formation (Scholpp *et al.,* 2007). The narrow ventral mid-diencephalic region of *shh* expression extends dorsally to the roof plate, and acts as the initial signal of the MDO. This extension of *shh* expression is possibly correlated with a decrease in dorsally -expressed retinoic acid (RA), which represses *shh* activity (Chambers *et al.*, 2007). The dorsoventral *shh* region of expression is bordered anteriorly by *Fez*, and posteriorly by *Irx* domains (Zeltser *et al.*, 2001). These two genes repress anteroposterior expansion of the *shh* domain (Jeong *et al.*, 2007). *Shh* signaling from the MDO is required for prethalamus and

thalamus development. Interestingly, regions anterior and posterior to the MDO respond in different ways to *shh* expression. It is thought that the earlier anterior Fez expression and posterior Irx expression cause these regions to be differentially primed, or have different competencies, and therefore respond in different ways to *shh* (Kiecker and Lumsden, 2004). The prethalamus lies anteriorly, and the thalamus posteriorly, to the MDO. The *shh* signaling induces the expression of *neurog1* in the thalamus and *ascl1a* in the prethalamus (Vue *et al.*, 2007). In fish, the expression of *ascl1a* is activated by her6, a protein that represses *neurog1* and is expressed in the prethalamus but not in the thalamus (Scholpp *et al.,* 2009). *Neurog1* is a marker for dopaminergic neural fate, and *ascl1a* is a marker for GABAergic interneuron fate. The thalamus and prethalamus are proliferative centers for dopaminergic neurons and GABAergic interneurons, respectively. Several genes are responsible for delineating the two separate developmental modules. Wnt proteins are implicated in thalamus development and are expressed in the MDO and thalamus, and are repressed by Lhx5 and Sfrp proteins within the prethalamus. Exogenous expression of wnts will transform prethalamus into thalamus, and experimental inhibition of wnts transforms thalamus into prethalamus (Braun *et al*., 2003). In zebrafish, Fgf8, Fgf3 and Zic2a are also necessary for prethalamic development, the latter two are positive regulators of *dlx1a/dlx2a* (Walshe and Mason, 2003; Sanek *et al.*, 2009). **Figure 18** shows a diagram of the GRN underlying GABAergic interneuron differentiation in the zebrafish diencephalon.

Figure 18: Diagram showing the gene regulatory network underlying GABAergic interneuron differentiation and migration in the zebrafish diencephalon. Two parralel *shh*- and *fez1*-mediated cascades lead to the activation of *ascl1a, dlx1a/dlx2a, dlx5a/dlx6a,* and *gad1a* expression. Thick horizontal lines indicate one gene (or node), in the network; gene names are listed under each line; thick bent arrows indicate active transcription of a given gene; thin lines with arrowheads indicate positive regulatory interactions between genes; thin lines ending with a short flat horizontal line indicate negative (inhibitory) regulatory interactions between genes; dashed thin lines suggest the presence of other unknown regulatory interactions; red colouring highlights genes and regulatory interactions investigated in this work; asterisks indicate genes whose orthologs play comparable regulatory roles in the mouse forebrain. (Information used from Jeong et al. 2007; Miyake et al. 2005; Pogoda et al. 2006; Sanek and Grinblat, 2008; Scholpp et al. 2006; Shinya et al. 2001; Varga et al. 2001; Walshe and Mason, 2003).


fez1
wnt1
pax2
wnt8b
lef1
zic2a
fgf3 / fgf8
shh
her6
ascl1a
Inhibition of
glutamatergic specification
neurog1
(I12b / URE2 enhancers)
dlx1a / dlx2a
(I56i / I56ii enhancers)
glutamic acid
decarboxylase
gad1a
dlx5a / dlx6a
* Orthologous mouse genes play comparable regulatory role

4.2.2 Down-regulation of *ascl1a* and *dlx1a/dlx2a* decreases *gad1a* expression in the zebrafish prethalamus

In summary, the coordinated expression of *shh, fez, her6, wnt* inhibitors, *fgf3* and *zic2a* is essential for prethalamic patterning, yields a regulatory environment leading to *ascl1a* and *dlx1a/dlx2a* expression, and determines prethalamic GABAergic interneuron fate specification. My work shows that *ascl1a* is a positive regulator of *dlx1a/dlx2a* and *dlx5a/dlx6a* in the ventral prethalamus. When *ascl1a* is down -regulated, *dlx1a/dlx2a* and *dlx5a/dlx6a* expression is diminished in this region. It is important to note that *ascl1a* expression, between 24 hpf and 48 hpf, is predominantly in the ventral and dorsal diencephalon, and is only minimally expressed in the telencephalon. This may explain why down –regulation of *ascl1a* only noticeably affects *dlx* expression in the diencephalon. When *dlx1a* and *dlx2a* are down -regulated together, an identical decrease in *gad1a* expression is seen in the prethalamus. In *dlx5a/dlx6a* double morphants there was no observed change in *gad1a* expression. It is possible that these genes lead to differentiation processes of GABAergic neurons which do not lead to the specific expression of the chosen marker, *gad1a.* Conversely, these results suggest that *dlx5a/dlx6a* do not play a role in the GRN underlying GABAergic fate specification. If this were the case, the expression of *dlx5a/dlx6a* may be important for cellular functions independent of GABAergic interneuron differentiation.

The similar prethalamic *gad1a* phenotype seen in both *ascl1a* and *dlx1a/dlx2a* morphants suggests either (1) that *ascl1a* independently positively regulates both *dlx* genes and *gad1a* in the prethalamus, or (2) *ascl1a* positively regulates *dlx1a/dlx2a* and *dlx5a/dlx6a,* but then *dlx1a/dlx2a* alone activate *gad1a* in the ventral prethalamus. If *dlx* and *gad1a* were separately regulated by *ascl1a,* we would not expect to see a decrease in *gad1a* expression in *dlx* morphants. This

supports the second regulatory scenario, in which *gad1a* is regulated by *ascl1a* through *dlx1a/dlx2a* activity. Further evidence for this is the partial rescue of the prethalamic *gad1a* phenotype *in ascl1a* morphants by exogenous expression of *dlx2a.* If *ascl1a* regulated *dlx1a/dlx2a* and *gad1a* through two different pathways, *dlx2a* exogenous expression would not be expected to rescue the *gad1a* phenotype. Since this is not the case, these data show *ascl1a* to be upstream of *dlx1a/dlx2a,* which is upstream of *gad1a.* Interestingly, although down-regulation of *dlx5a/dlx6a* does not noticeably alter *gad1a* expression, exogenous expression of *dlx5a* in *ascl1a* morphants has a similar capacity as *dlx2a* in rescuing the *gad1a* phenotype. To explain this, it must be iterated that exogenous expression of a gene, by injecting mRNA into an embryo, is not a precise means to experimentally replicate the endogenous activity of a gene. In the case of a highly conserved family of homeodomain transcription factors such as the dlx proteins, each protein is likely to have very similar biochemical properties. Their spatiotemporally unique expression patterns and various regulatory roles during development may be more contingent on differences in regulatory regions than in the biochemical properties of the proteins. So, although dlx5a and dlx6a may not endogenously contribute to the activation of *gad1a* expression in the prethalamus, the artificially over-expressed dlx5a protein could perform a similar regulatory role to dlx1a and dlx2a, thus rescuing the *gad1a* phenotype in *ascl1a* morphants. It is also important to note that in both *ascl1a* and *dlx1a/dlx2a* morphants, the expression of *gad1a* is only partially downregulated in the diencephalon. This strongly suggests that there are other important regulators of diencephalic *gad1a* expression at this developmental stage, which have not been addressed in this work.

The observation that forebrain *dlx5a/dlx6a* expression is not noticeably altered in *dlx1a/dlx2a* double morphants may be interpreted in at least two ways. (1) It is possible that

Dlx1a and Dlx2a proteins to do not play an important role in regulating *dlx5a/dlx6a* expression in the forebrain. Hence, down-regulation of the former pair has no impact on the expression of the latter. (2) Conversely, the Dlx1a and Dlx2a proteins do regulate *dlx5a/dlx6a* expression, but in a cross-regulatory manner that induces auto-upregulation of *dlx5a/dlx6a* when Dlx1a/Dlx2a activators are absent. Down-regulation of *dlx1a/dlx2a* in transgenic lines expressing reporter genes activated by intergenic enhancers leads to decreased reporter gene expression. This indicates that the Dlx1a/Dlx2a proteins are capable of binding enhancers that play important roles in *dlx* forebrain expression. Preliminary work is showing that mice lacking the I56i enhancer exhibit decreased *Dlx5* expression, but a simultaneous slight up-regulation of *Dlx1* (personal communication from Crystal Esau). These data suggests that *Dlx2* is up -regulated to compensate for the decrease in *Dlx5* expression. A similar compensatory effect may be at work in the zebrafish *dlx1a/dlx2a* morphants.

4.2.3 Evolutionary perspectives on developmental GRNs in the vertebrate forebrain

How do developmental GRNs diverge from one another between lineages? Is there any predictability in which components of a GRN are likely to change over evolutionary time? There are two concepts to address before making evolutionary deductions of my work. Firstly, the likelihood of major evolutionary changes occurring in GRNs is related to the temporal dimension of organismal development. To generalize, GRNs involved in processes that occur later in development, (e.g. the terminal differentiation of cell types, or the formation of superficial microstructures), are more susceptible to faster evolutionary change than those involved in earlier, and more fundamental body / tissue patterning (Davidson, 2006). The idea here is that changes in early morphogenesis are likely to have more severe downstream impacts than those in

later development, and are therefore more likely to be lost in evolution. An elegant corollary to this is that deep evolutionary morphological divergence (separating phyla, classes and orders) arises through changes in *early* developmental GRNs; bringing major differences in fundamental body plan organization (Davidson, 2006). Conversely, 'shallow' evolutionary morphological divergence (separating families, genus and species) is caused by novelties in GRNs operating *later* in development. In the present work I made comparisons between vertebrate lineages that diverged over 300 mya and belong to different classes. I looked at GRNs active during early development of the forebrain. According to the above evolutionary paradigm, it is difficult to predict what will be found when studying early developmental GRNs in deeply divergent organisms.

The second concept pertaining to GRN evolution is that the genetic constituents, or nodes of a GRN have varying likelihoods of undergoing functional change. When examining the architecture of a GRN there are inevitably some nodes that receive more regulatory inputs, and produce a greater number of regulatory outputs, than others. The higher interconnectedness of these nodes relates to a more central role in the overall GRN activity. It has therefore been hypothesized that when comparing analogous GRNs in different species, discrepancies should first be found in the activity of less interconnected GRN components (Fischer and Smith, 2012).

My work has shown that divergence in the activity of GRNs responsible for forebrain neurodevelopment has occurred between mice and zebrafish vertebrate lineages. Specifically, *Dlx* gene expression contributes to the correct timing of migration and differentiation of GABAergic interneuron precursors in the ventral telencephalon and diencephalon in mice. In zebrafish, although showing highly similar *dlx* expression patterns in the forebrain, these genes

are only implicated in diencephalic GABAergic neurogenesis. Unlike in mice, *dlx5a* and *dlx6a* do not play a detectable regulatory role during this process in zebrafish.

Among studied extant vertebrates, the early patterning of the diencephalon has been highly conserved. The roles of *Shh, Fez, Otx, Wnts, Ascl1, Neurog1* and *Fgfs* in prethalamic and thalamic regionalization are similar between mouse, chick, Xenopus and zebrafish (Scholpp and Lumsden, 2010). Prethalamic specification of GABAergic interneuron fate begins with the expression of the proneural gene *Ascl* (or *Mash*). As these neural precursors exit the cell cycle and begin to migrate / differentiate, neural proteins like Dlx are expressed and activate other downstream genes specific to this cellular lineage, including *Gad*. It is evident that the regulatory relationship between *ascl1a*, *dlx1a/dlx2a* and *gad1a* is present in the developing zebrafish prethalamus. It is possible that in the last common ancestor to mice and zebrafish *Dlx5* and *Dlx6* were also involved in this developmental process, but this role was gradually lost in the zebrafish lineage. If this were the case, these genes may have been examples of nodes lacking interconnectedness in a GRN and were more vulnerable to exclusion. Conversely, if in the last common ancestor *Dlx5* and *Dlx6* were not involved in this process, they may have been co-opted for this function in the tetrapod lineage.

One interesting aspect of these findings is that the conservation of the diencephalic GRN regulating GABAergic interneuron development, and the apparent telencephalic divergence of the same process, reinforces the modular model of evolution. Segmentation is integral to organismal development. First broad and rudimentary during embryogenesis, and later specific and precise during maturation, segmentation allows for different parts of the body to develop relatively independently and according to localized genetic programs. This phenomenon makes it possible for evolution to 'tweak' a certain part of the body, by altering the genetic basis for its

development, with minimal changes to nearby structures. The brain is no exception here, and is thought to develop with neuromeric zones that regionalize the brain anteroposteriorly and dorsoventrally. There is a wide diversity in brain morphology among extant vertebrates, and major defining differences between them are relative changes in size of discrete brain regions. The basic brain plan is highly conserved, but there are lineage –specific variations in how certain segments develop. For example, mammals have a relatively large cerebral cortex (telencephalon derivative), teleosts have a relatively large optic tectum (diencephalon derivative), and cartilaginous fish have a relatively large olfactory bulb (telencephalon derivative). As evolution increases the size of a given segment of the brain, changes in developmental GRNs must necessarily occur. The increased proliferation of neurons / glia and extended growth of vascular structures (among many other processes), is dependent on modifications in signaling and patterning cascades. Therefore, when comparing developmental GRN architecture in the brains of two species one would expect conservation in highly similar, and divergence in highly dissimilar, brain regions.

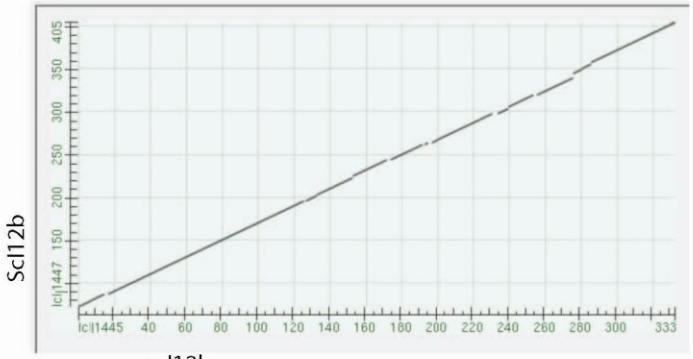
In the present work, the main differences seen between zebrafish and mouse forebrain developmental GRNs is in the telencephalon. It is an interesting possibility that these telencephalic changes in *Dlx* gene expression contributed to the increased cortex size found in the lineage leading to mammals. It is simplistic to causally relate the increased volume of brain regions with the gain of specific functions or behaviours. Instead, lineage –specific variations in how neurons are organized in the brain, and the manner in which they are interconnected, can equally contribute to novelties in brain function and animal behaviour. Nonetheless, these mechanisms producing novel brain phenotypes are equally dependent on changes in early developmental GRN dynamics. Disparities in neural migration paths, differential growth of

axons and dendrites, and the reorganization of neural circuit construction, can all give rise to new phenotypes that are subject to natural selection. As 20% of mammalian cortical neurons are GABAergic interneurons, it is likely that with increased overall cortical volume came an increase in GABAergic interneuron proliferation and migration during embryogenesis. It is possible that as telencephalic GRN architecture changed between the lineages leading to zebrafish and mice, the developmental roles played by *Dlx* transcription factors in specifying GABAergic interneurons also diverged.

Appendix A: BLAST analysis of I12b and I56i orthologous enhancer sequences of zebrafish (drI12b, drI56i), dogfish (scI12b, scI56i) and mouse (mI12b, mI56i). Dot matrix (top) and shown sequence comparison (bottom) between (A) mI12b and scI12b, which share 78% identity, (B) drI12b and scI12b, which share 78% identity, (C) mI12b and drI12b, 73% identity, (D) mI56i and drI56i, 86%, (E) mI56i and scI56i, 79% identity, (F) drI56i and scI56i, 78% identity. Data gathered and analyzed using the Basic Local Alignment Sequence Tool (BLAST) found at http://blast.ncbi.nlm.nih.gov/.

(A)

78% identity



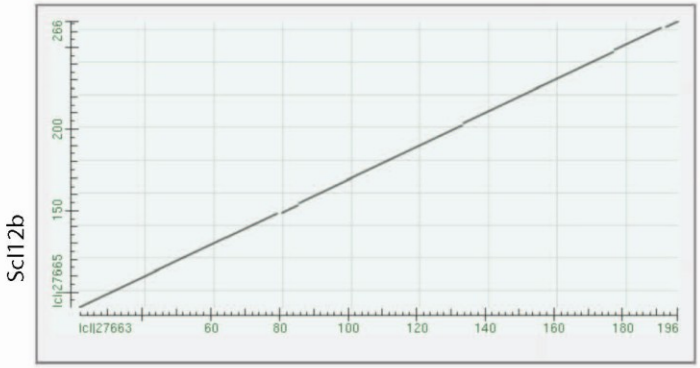
Score	Expect	Identities	Gaps	Strand
242 bits(268)	2e-68	269/346(78%)	26/346(7%)	Plus/Plus

ml12b
sc l12b

```
Query  1    CGCGTACAGCTGCAAACCCAAGAGGGTCAGCATCATTTCACTGTATTCTCTTCTTGATTA  60
Sbjct  73   CGCATACAGCTGCAA--CCAAGAGGCTCGGCATTATTACTCTGTATTCTCTACTTGATTA  130

Query  61   CAAGCCGGGCCCATCAAACACAACATAATTACAGTAATTTCAGGTTTATTTATTCTAATG  120
Sbjct  131  CAAGCCGAGGCCATCAAACGGACTATAATTACAGTAATTTTGAGTTTATTTATTCTAATG  190

Query  121  CAGTTTCCCCATC-TCTCTGGTAATTATGAGCAA--TTTTTTCGCCCAGGGAATCTTTTT  177
Sbjct  191  TAGTTT-CGTATCATTTGGGGTAATTATGAGCAATTTTTTTTCCGCCGGGGAATC--TTT  247

Query  178  GCATTAACAAAAGAGATAACGCACTGAAAGCCAAATTTGCTGCGCATTGAGAAAAGGGaa  237
Sbjct  248  GCATTAACAAAAGAG-TA--GCAGTGAAAGCCACATTTGAATAGCATTGAGAAA---GAA  301

Query  238  aaa--aaaaTCAAATAGGTGCGAGCTGCCATCTCTGCAATT-----CTCCG-GTACC--G  287
Sbjct  302  GAACCAAACTAAAATAGGT--GTGCTTCCATCTTTGAAATTCAGAACTAGGTTTATCATG  359

Query  288  GAGCCCGCAAATTGCTTGCAGGTGTATGGAGCAAGCTTGTCAATGG  333
Sbjct  360  AGGTCTGCAAATTGCTTGCAGGTGTATGGAGGAGCATTTTCAATGG  405
```

(B)

78% identity



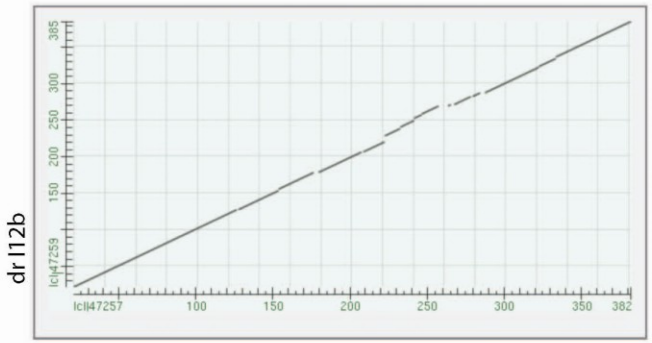
Score	Expect	Identities	Gaps	Strand
127 bits(140)	9e-34	139/178(78%)	5/178(2%)	Plus/Plus

drl12b
sc l12b

```
Query  22   AGAGGGTCAGCGTTATTTCACTCTATTCTCCTCTTGATTACGCGCCGGGCCCATCAAACC  81
Sbjct  91   AGAGGCTCGGCATTATTACTCTGTATTCTCTACTTGATTACAAGCCGAGGCCATCAAA-C  149

Query  82   AGAC-ATAATTACAGTCATTCCCTCCTTACTTATTCTAATGCAGCTTCCCATC-TACGGG  139
Sbjct  150  GGACTATAATTACAGTAATTTTGAGTTTATTTATTCTAATGTAGTTTCGTATCATTTGGG  209

Query  140  ATAATTATGAGCAAtttttttCGCACGTGGAATCTTT-TGTTAACAAAAGAGATAGC  196
Sbjct  210  GTAATTATGAGCAATTTTTTTTCCGCCGGGGAATCTTTGCATTAACAAAAGAG-TAGC  266
```

(C)

73% identity



ml12b

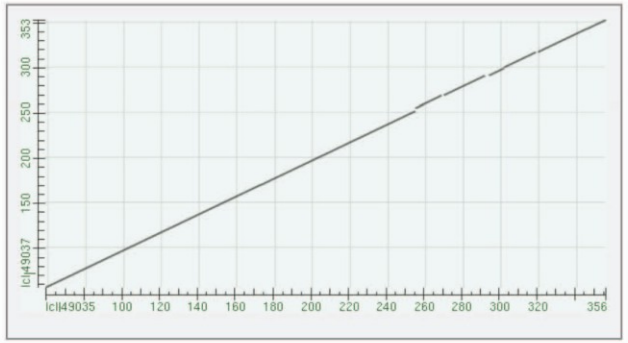
ml12b
dr l12b

Score	Expect	Identities	Gaps	Strand
192 bits(212)	2e-53	278/381(73%)	36/381(9%)	Plus/Plus

```
Query  21   AGAGGGTCAGCATCATTTCACTGTATTCTCTTCTTGATTACAAGCCGGGCCCATCAAACA  80
Sbjct  22   AGAGGGTCAGCGTTATTTCACTCTATTCTCCTCTTGATTACGCGCCGGGCCCATCAAACC  81
Query  81   CAACATAATTACAGTAATTTCAGGTTTATTTATTCTAATGCAGTTTCCCCATCTCTCTGG  140
Sbjct  82   AGACATAATTACAGTCATTCCCTCCTTACTTATTCTAATGCAGCTT-CCCATCTACGGGA  140
Query  141  TAATTATGAGCAA--TTTTTTCGCCCAGGGAATCTTTTTGCATTAACAAAAGAGATAACG  198
Sbjct  141  TAATTATGAGCAATTTTTTTTTCGCACGTGGAATCTTTT---GTTAACAAAAGAGATAGCG  197
Query  199  CACTGAAAGCCAAATTTGCTGCGC--------ATTGAGAAAA--GGGaaaaaa---aaaT  245
Sbjct  198  CGCTGAAAG-AAAATTTGCTGTGCGACGGGGGAAAAAGAAAAGGGGGAAAAAAGAGAGAG  256
Query  246  C-AAATAGGTGCGAGCTGCCATCTCTGCAATTCTCCGGTACCGGAGCCCGCAAATTGCTT  304
Sbjct  257  CGAAATAGGTGCG------CA--GCGGTAGTTAT-AAGTA---AGGCTGGCAAATTGCTT  304
Query  305  GCAGGTGTATGGAGCAAG-CTTGTCAATGG--CCGGGCCTCCAAATTAGCAAATGCACAG  361
Sbjct  305  GCAGGTGTATACCGAGAGATTTGTCAATGGAAAGAGGGATCAAAATGAGCAAAAGCACAA  364
Query  362  CAGCGAAGTAATGAAGACAGA  382
Sbjct  365  TGTCGGAGTGATGAAGGCAGA  385
```

(D)

86% identity



ml56i

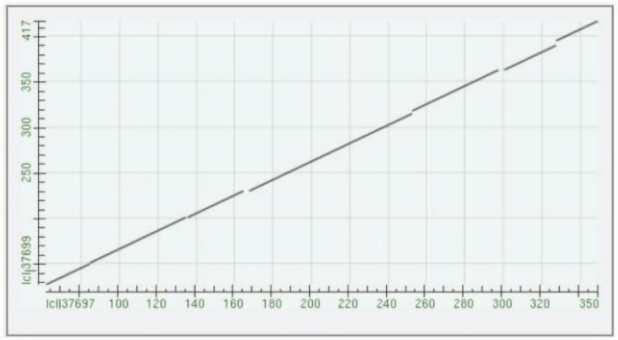
ml56i
dr l56i

Score	Expect	Identities	Gaps	Strand
127 bits(140)	9e-34	139/178(78%)	5/178(2%)	Plus/Plus

```
Query  22   AGAGGGTCAGCGTTATTTCACTCTATTCTCCTCTTGATTACGCGCCGGGCCCATCAAACC  81
Sbjct  91   AGAGGCTCGGCATTATTACTCTGTATTCTCTACTTGATTACAAGCCGAGGCCATCAAA-C  149
Query  82   AGAC-ATAATTACAGTCATTCCCTCCTTACTTATTCTAATGCAGCTTCCCATC-TACGGG  139
Sbjct  150  GGACTATAATTACAGTAATTTTGAGTTTATTTATTCTAATGTAGTTTCGTATCATTTGGG  209
Query  140  ATAATTATGAGCAAtttttttttCGCACGTGGAATCTTT-TGTTAACAAAAGAGATAGC  196
Sbjct  210  GTAATTATGAGCAATTTTTTTTTCCGCCGGGGAATCTTTGCATTAACAAAAGAG-TAGC  266
```

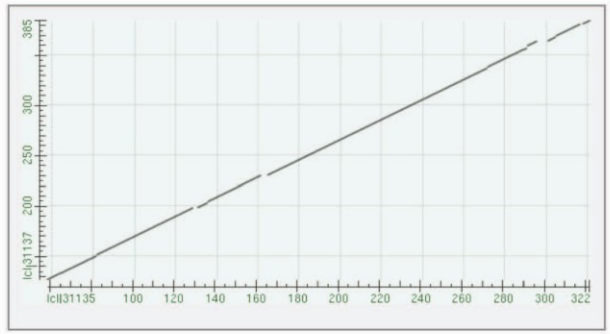
(E)

79% identity



(F)

78% identity



ml56i
scl56i

Range 128 to 417: Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
244 bits(270)	4e-69	235/297(79%)	16/297(5%)	Plus/Plus

```
Query  63   CAGGATCAATTCTGAACAAAGCC-TCCAGCTGCAGTGCCATCCAATTTGAAGCAGACATT  121
            || |||||||  ||  |||||||  ||  ||||||||||   ||||||||||  ||||  ||  |||
Sbjct  128  CACGATCAATACTAAACAAAACCCTCCAGCTGCAGACCCATCCAATTTAAAGCTGATATT  187

Query  122  GGGGACAATTTAAGGTTTTTATCCACAAGAAGGttttttttCCATTCTCTTAAATGCAGCC  181
             ||  |||||||  |  ||  ||||  |||  ||||  |   ||    ||||    | ||||  ||  ||
Sbjct  188  TGGTACAATTTGCG-TTATTATTCACGAGAAAGGCTTAGACCAT---CATAAAAGCGGCT  243

Query  182  ATAATTAGAGTAATTTTTCATGTAGCCCGCTGATTACAGCGTTTTTACCGTCAAAGATAA  241
             ||||||||  |||||||||   |||  ||||||||||||||||||||||||  |||||||  |||
Sbjct  244  ATAATTAGGGTAATTTTTATTGTGGCCCGCTGATTACAGCGTTTTTACAGTCAAAGGTAA  303

Query  242  TTACCTGTAATT---TTCCACTTTTAATACTAAAAAGCCATCTTTATTTAGATTCAGGAA  298
             ||||  |||||||    ||||||||||||||||||  |||  ||||||||||  |||||||||
Sbjct  304  TTACATGTAATTTTCTTCCACTTTTAATACTAAAGAGCTATCTTTATTTAAATTCAGGAA  363

Query  299  CAGGAAAGGCGAAACAAAAGAGGGAAATTA-----TTCTGTTATTCATACACAAATT  350
                |||  |  ||||||||||||  |  |  ||     ||  ||||||||||  ||  ||||
Sbjct  364  ---GAAGAGTGAAACAAAAGAGAAAGAATAGAACGTTGTGTTATTCATAAACCAATT  417
```

drl56i
sc l56i

Score	Expect	Identities	Gaps	Strand
196 bits(216)	2e-54	211/270(78%)	17/270(6%)	Plus/Plus

```
Query  59   ACAGGATCGATCCTGAACAAAGCA-TCCAGCTGCAGTCTCGTTCAATATGAAGGAGATAT  117
            |||  ||||  ||  ||  ||||||  |  ||||||||||  |  |  ||||  |  |||  |||||
Sbjct  127  ACACGATCAATACTAAACAAAACCCTCCAGCTGCAGACCCATCCAATTTAAAGCTGATAT  186

Query  118  TTGGGACAATTTATGGGTT-TTATCCAAGAGAGGGttttttttCCATTCTCATAAATGCAG  176
             ||||  |||||||  |  |||  ||||  ||  ||||  |  ||  ||||  ||||||  ||  |
Sbjct  187  TTGGTACAATTT--GCGTTATTATTCACGAGAAAGGCTTAGACCAT---CATAAAAGCGG  241

Query  177  ACATAATTAGGGTAATTTTTGATGTAGCCCGCTGATTACAGCGTTTTTACCGTCAAAGAT  236
             ||||||||||||||||||  |||  ||||||||||||||||||||||||||  |||||||  |
Sbjct  242  CTATAATTAGGGTAATTTTTATTGTGGCCCGCTGATTACAGCGTTTTTACAGTCAAAGGT  301

Query  237  AATTACCTGTAATTTTCTACCACTTTTAATACTAAA-AGGCATCTTTATTTGGATT--TG  293
             ||||||  ||||||||||  ||||||||||||||||  ||  ||||||||||  |||  |
Sbjct  302  AATTACATGTAATTTTCTTCCACTTTTAATACTAAAGAGCTATCTTTATTTAAATTCAGG  361

Query  294  AAGTGGCAAAGA-CGAAACAAAAGATGAAA  322
             |||  ||||  ||||||||||  ||||
Sbjct  362  AAG-----AAGAGTGAAACAAAAGA-GAAA  385
```

REFERENCES

Anderson, S. A., Qiu, M., Bulfone, A., Eisenstat, D. D., Meneses, J., Pedersen, R. and Rubenstein, J. L. R. (1997a). Mutations of the Homeobox Genes *Dlx-1* and *Dlx-2* Disrupt the Striatal Subventricular Zone and Differentiation of Late-Born Striatal Neurons. *Neuron* 19, 27-37.

Ayala, R., Shu, T. and Tsai, L. H. (2007). Trekking Across the Brain: The Journey of Neuronal Migration. *Cell* 128, 29-43.

Babu, M. M., Luscombe, N. M., Aravind, L., Gerstein, M., & Teichmann, S. A. (2004). Structure and evolution of transcriptional regulatory networks. *Current Opinion in Structural Biology, 14*, 283-291.

Benes, F. M., & Berretta, S. (2001). GABAergic interneurons: Implications for understanding schizophrenia and bipolar disorder. *Neuropsychopharmacology, 25*, 1-27.

Bond, A. M., Vangompel, M. J., Sametsky, E. A., Clark, M. F., Savage, J. C., Disterhoft, J. F. and Kohtz, J. D. (2009). Balanced Gene Regulation by an Embryonic Brain ncRNA is Critical for Adult Hippocampal GABA Circuitry. *Nat. Neurosci.* 12, 1020-1027.

Braun, M. M., Etheridge, A., Bernard, A., Robertson, C. P. and Roelink, H. (2003). Wnt Signaling is Required at Distinct Stages of Development for the Induction of the Posterior Forebrain. *Development* 130, 5579-5587.

Chambers, D., Wilson, L., Maden, M., & Lumsden, A. (2007). RALDH-independent generation of retinoic acid during vertebrate embryogenesis by CYP1B1. *Development, 134*, 1369-1383.

Cohen, S. M., Bronner, G., Kuttner, F., Jurgens, G. and Jäckle, H. (1989). *Distal-Less* Encodes a Homeodomain Protein Required for Limb Development in *Drosophila*. *Nature* 338, 432-434.

Cohen, S. M. and Jurgens, G. (1989). Proximal-Distal Pattern Formation in Drosophila: Cell Autonomous Requirement for Distal-Less Gene Activity in Limb Development. *EMBO J.* 8, 2045-2055.

Colasante, G., Collombat, P., Raimondi, V., Bonanomi, D., Ferrai, C., Maira, M., Yoshikawa, K., Mansouri, A., Valtorta, F., Rubenstein, J. L. et al. (2008). Arx is a Direct Target of Dlx2 and Thereby Contributes to the Tangential Migration of GABAergic Interneurons. *J. Neurosci.* 28, 10674-10686.

Davidson, E. H. (2006). *Gene regulatory networks in development and evolution*. New York: Elsevier. Chapters 1-3.

Dickmeis, T. and Muller, F. (2005). The Identification and Functional Characterisation of Conserved Regulatory Elements in Developmental Genes. *Brief Funct. Genomic Proteomic* 3, 332-350.

Dollé, P., Price, M. and Duboule, D. (1992). Expression of the Murine *Dlx-1* Homeobox Gene during Facial, Ocular and Limb Development. *Differentiation* 49, 93-99.

Ebbesson, S. O. E. (1980). The parcellation theory and its relation to interspecific variability in brain organization, evolutionary and ontogenetic development, and neuronal plasticity. *Cell and Tissue Research, 213*, 179-212.

Feledy, J. A., Morasso, M. I., Jang, S. -. and Sargent, T. D. (1999). Transcriptional Activation by the Homeodomain Protein Distal-Less 3. *Nucl. Acids Res.* 27, 764-770.

Feng, J., Bi, C., Clark, B. S., Mady, R., Shah, P. and Kohtz, J. D. (2006). The Evf-2 Noncoding RNA is Transcribed from the Dlx-5/6 Ultraconserved Region and Functions as a Dlx-2 Transcriptional Coactivator. *Genes Dev.* 20, 1470-1484.

Fischer, A. H. L., & Smith, J. (2012). Evo–Devo in the era of gene regulatory networks. *Integrative and Comparative Biology* 52, 842-849.

Gelman, D. M., Martini, F. J., Nobrega-Pereira, S., Pierani, A., Kessaris, N. and Marin, O. (2009). The Embryonic Preoptic Area is a Novel Source of Cortical GABAergic Interneurons. *J. Neurosci.* 29, 9380-9389.

Ghanem, N., Jarinova, O., Amores, A., Hatch, G., Park, B. K., Rubenstein, J. L. R. and Ekker, M. (2003). Regulatory Roles of Conserved Intergenic Domains in Vertebrate *Dlx* bigene Clusters. *Genome Res.* 13, 533-543.

Ghanem, N., Yu, M., Poitras, L., Rubenstein, J. L. and 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. *Dev. Biol.* 322, 415-424.

Guillemot, F. (2007). Spatial and Temporal Specification of Neural Fates by Transcription Factor Codes. *Development* 134, 3771-3780.

Guinazu, M. F., Chambers, D., Lumsden, A., & Kiecker, C. (2007). Tissue interactions in the developing chick diencephalon. *Neural Dev, 2*, 25-40.

Hauptmann, G., & Gerster, T. (2000). Regulatory gene expression patterns reveal transverse and longitudinal subdivisions of the embryonic zebrafish forebrain. *Mechanisms of Development,* 91,105-118.

Heisenberg, C. P., Houart, C., Take-uchi, M., Rauch, G. J., Young, N., Coutinho, P., Masai, I., Caneparo, L., Concha, M.L., Geisler, R., Dale, T.C., Wilson, S.W., Stemple, D.L., and Geisler, R. (2001). A mutation in the Gsk3–binding domain of zebrafish Masterblind/Axin1 leads to a fate transformation of telencephalon and eyes to diencephalon. *Genes & Development, 15*(11), 1427-1434.

Houart, C., Caneparo, L., Heisenberg, C., Barth, K., Take-Uchi, M. and Wilson, S. (2002). Establishment of the Telencephalon during Gastrulation by Local Antagonism of Wnt Signaling. *Neuron* 35, 255-265.

Inverardi, F., Beolchi, M. S., Ortino, B., Moroni, R. F., Regondi, M. C., Amadeo, A. and Frassoni, C. (2007). GABA Immunoreactivity in the Developing Rat Thalamus and Otx2 Homeoprotein Expression in Migrating Neurons. *Brain Res. Bull.* 73, 64-74.

Jeong, J. Y., Einhorn, Z., Mathur, P., Chen, L., Lee, S., Kawakami, K. and Guo, S. (2007). Patterning the Zebrafish Diencephalon by the Conserved Zinc-Finger Protein Fezl. *Development* 134, 127-136.

Kiecker, C. and Lumsden, A. (2004). Hedgehog Signaling from the ZLI Regulates Diencephalic Regional Identity. *Nat. Neurosci.* 7, 1242-1249.

Kim, A. S., Lowenstein, D. H., & Pleasure, S. J. (2001). Wnt receptors and wnt inhibitors are expressed in gradients in the developing telencephalon. *Mechanisms of Development,* 103, 167-172.

Kimura, M. (1983). Rare Variant Alleles in the Light of the Neutral Theory. *Mol. Biol. Evol.* 1, 84-93.

Kudoh, T., Wilson, S. W. and Dawid, I. B. (2002). Distinct Roles for Fgf, Wnt and Retinoic Acid in Posteriorizing the Neural Ectoderm. *Development* 129, 4335-4346.

Kulkarni, M. M. and Arnosti, D. N. (2003). Information Display by Transcriptional Enhancers. *Development* 130, 6569-6575.

Kuwajima, T., Nishimura, I., & Yoshikawa, K. (2006). Necdin promotes GABAergic neuron differentiation in cooperation with dlx homeodomain proteins. *The Journal of Neuroscience,* 26, 5383-5392.

Le, T. N., Du, G., Fonseca, M., Zhou, Q. P., Wigle, J. T. and Eisenstat, D. D. (2007). Dlx Homeobox Genes Promote Cortical Interneuron Migration from the Basal Forebrain by Direct Repression of the Semaphorin Receptor Neuropilin-2. *J. Biol. Chem.* 282, 19071-19081.

Liu, J. K., Ghattas, I., Liu, S., Chen, S. and Rubenstein, J. L. R. (1997). Dlx genes Encode DNA-Binding Proteins that are Expressed in an Overlapping and Sequential Pattern during Basal Ganglia Differentiation. *Dev. Dyn*. 210, 498-512.

MacDonald, R. B., Debiais-Thibaud, M., Talbot, J. C., & Ekker, M. (2010). The relationship between dlx and gad1 expression indicates highly conserved genetic pathways in the zebrafish forebrain. *Developmental Dynamics,* 239, 2298-2306.

Martin, S. C., Heinrich, G. and Sandell, J. H. (1998). Sequence and Expression of Glutamic Acid Decarboxylase Isoforms in the Developing Zebrafish. *J. Comp. Neurol.* 396, 253-266.

Masuda, Y., Sasaki, A., Shibuya, H., Ueno, N., Ikeda, K. and Watanabe, K. (2001). Dlxin-1, a Novel Protein that Binds Dlx5 and Regulates its Transcriptional Function. *J. Biol. Chem.* 276, 5331-5338.

Marin, O., Anderson, S. A. and Rubenstein, J. L. (2000). Origin and Molecular Specification of Striatal Interneurons. *J. Neurosci.* 20, 6063-6076.

Miyake, A., Nakayama, Y., Konishi, M. and Itoh, N. (2005). Fgf19 Regulated by Hh Signaling is Required for Zebrafish Forebrain Development. *Dev. Biol.* 288, 259-275.

Mueller, T. and Wullimann, M. F. (2002). BrdU-, neuroD (Nrd)- and Hu-Studies Reveal Unusual Non-Ventricular Neurogenesis in the Postembryonic Zebrafish Forebrain. *Mech. Dev.* 117, 123-135.

Mueller, T. and Wullimann, M. F. (2003). Anatomy of Neurogenesis in the Early Zebrafish Brain. *Brain Res. Dev. Brain Res.* 140, 137-155.

Mueller, T., Vernier, P. and Wullimann, M. F. (2006). A Phylotypic Stage in Vertebrate Brain Development: GABA Cell Patterns in Zebrafish Compared with Mouse. *J. Comp. Neurol.* 494, 620-634.

Nieuwenhuys, R. and Meek, J. (1990). The Telencephalon of Actinopterygian Fishes. In *Comparative Structure and Evolution of the Cerebral Cortex* (ed. E. G. Jones and A. Peters), pp. 31-73. New York: Plenum.

Northcutt, R. G. (2002). Understanding vertebrate brain evolution. *Integrative and Comparative Biology, 42*(4), 743-756.

Ortino, B., Inverardi, F., Morante-Oria, J., Fairen, A. and Frassoni, C. (2003). Substrates and Routes of Migration of Early Generated Neurons in the Developing Rat Thalamus. *Eur. J. Neurosci.* 18, 323-332.

Osorio, J., Mueller, T., Retaux, S., Vernier, P. and Wullimann, M. F. (2010). Phylotypic Expression of the bHLH Genes Neurogenin2, Neurod, and Mash1 in the Mouse Embryonic Forebrain. *J. Comp. Neurol.* 518, 851-871.

Panganiban, G. and Rubenstein, J. L. R. (2002). Developmental Functions of the *Distal-less/Dlx* homeobox Genes. *Development* 129, 4371-4386.

Panne, D., Maniatis, T. and Harrison, S. C. (2007). An Atomic Model of the Interferon-Beta Enhanceosome. *Cell* 129, 1111-1123.

Park, B. K., Sperber, S. M., Choudhury, A., Ghanem, N., Hatch, G. T., Sharpe, P. T., Thomas, B.L., 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.

Pennacchio, L. A., Ahituv, N., Moses, A. M., Prabhakar, S., Nobrega, M. A., Shoukry, M., Minovitsky, S., Dubchak, I., Holt, A., Lewis, K. D. et al. (2006). In Vivo Enhancer Analysis of Human Conserved Non-Coding Sequences. *Nature* 444, 499-502.

Petryniak, M. A., Potter, G. B., Rowitch, D. H. and Rubenstein, J. L. (2007). Dlx1 and Dlx2 Control Neuronal Versus Oligodendroglial Cell Fate Acquisition in the Developing Forebrain. *Neuron* 55, 417-433.

Pogoda, H. M., von der Hardt, S., Herzog, W., Kramer, C., Schwarz, H. and Hammerschmidt, M. (2006). The Proneural Gene ascl1a is Required for Endocrine Differentiation and Cell Survival in the Zebrafish Adenohypophysis. *Development* 133, 1079-1089.

Poitras, L., Ghanem, N., Hatch, G. and Ekker, M. (2007). The Proneural Determinant MASH1 Regulates Forebrain Dlx1/2 Expression through the I12b Intergenic Enhancer. *Development* 134, 1755-1765.

Potter, G. B., Petryniak, M. A., Shevchenko, E., McKinsey, G. L., Ekker, M. and Rubenstein, J. L. (2009). Generation of Cre-Transgenic Mice using Dlx1/Dlx2 Enhancers and their Characterization in GABAergic Interneurons. *Mol. Cell. Neurosci.* 40, 167-186.

Scholpp, S., Foucher, I., Staudt, N., Peukert, D., Lumsden, A. and Houart, C. (2007). Otx1l, Otx2 and Irx1b Establish and Position the ZLI in the Diencephalon. *Development* 134, 3167-3176.

Venkatesh, B., Kirkness, E. F., Loh, Y. H., Halpern, A. L., Lee, A. P., Johnson, J., Dandona, N.,

Viswanathan, L. D., Tay, A., Venter, J. C. et al. (2007). Survey Sequencing and Comparative Analysis of the Elephant Shark (Callorhinchus Milii) Genome. *PLoS Biol.* 5, e101.

Rakic, P. (1972). Mode of Cell Migration to the Superficial Layers of Fetal Monkey Neocortex. *J. Comp. Neurol.* 145, 61-83.

Roth, G., & Wullimann, M. F. (2001). *Brain evolution and cognition* Wiley-Spektrum.

Sanek, N. A. and Grinblat, Y. (2008). A Novel Role for Zebrafish zic2a during Forebrain Development. *Dev. Biol.* 317, 325-335.

Sanek, N. A., Taylor, A. A., Nyholm, M. K. and Grinblat, Y. (2009). Zebrafish zic2a Patterns the Forebrain through Modulation of Hedgehog-Activated Gene Expression. *Development* 136, 3791-3800.

Scholpp, S., Delogu, A., Gilthorpe, J., Peukert, D., Schindler, S. and Lumsden, A. (2009). Her6 Regulates the Neurogenetic Gradient and Neuronal Identity in the Thalamus. *Proc. Natl. Acad. Sci. U. S. A.* 106, 19895-19900.

Scholpp, S., and Lumsden, A. (2010). Building a bridal chamber: Development of the thalamus. *Trends in Neurosciences, 33*(8), 373-380.

Scholpp, S., Wolf, O., Brand, M. and Lumsden, A. (2006). Hedgehog Signalling from the Zona Limitans Intrathalamica Orchestrates Patterning of the Zebrafish Diencephalon. *Development* 133, 855-864.

Shinya, M., Koshida, S., Sawada, A., Kuroiwa, A. and Takeda, H. (2001). Fgf Signalling through MAPK Cascade is Required for Development of the Subpallial Telencephalon in Zebrafish Embryos. *Development* 128, 4153-4164.

Shirasaki, R. and Pfaff, S. L. (2002). Transcriptional Codes and the Control of Neuronal Identity. *Annu. Rev. Neurosci.* 25, 251-281.

Sperber, S. M., Saxena, V., Hatch, G. and Ekker, M. (2008). Zebrafish dlx2a Contributes to Hindbrain Neural Crest Survival, is Necessary for Differentiation of Sensory Ganglia and Functions with dlx1a in Maturation of the Arch Cartilage Elements. *Dev. Biol.* 314, 59-70.

Stuhmer, T., Anderson, S. A., Ekker, M. and Rubenstein, J. L. R. (2002a). Ectopic Expression of the *Dlx* genes Induces Glutamic Acid Decarboxylase and *Dlx* expression. *Development* 129, 245-252.

Stuhmer, T., Puelles, L., Ekker, M. and Rubenstein, J. L. (2002b). Expression from a Dlx Gene Enhancer Marks Adult Mouse Cortical GABAergic Neurons. *Cereb. Cortex* 12, 75-85.

Thomas, B. L., Liu, J. K., Rubenstein, J. L. R. and Sharpe, P. T. (2000). Independent Regulation of *Dlx2* expression in the Epithelium and Mesenchyme of the First Branchial Arch. *Development* 127, 217-224.

Varga, Z. M., Amores, A., Lewis, K. E., Yan, Y. L., Postlethwait, J. H., Eisen, J. S. and Westerfield, M. (2001). Zebrafish Smoothened Functions in Ventral Neural Tube Specification and Axon Tract Formation. *Development* 128, 3497-3509.

Vieira, C., Garda, A. L., Shimamura, K., & Martinez, S. (2005). Thalamic development induced by *Shh* in the chick embryo. *Developmental Biology,* 284, 351-363.

Vue, T. Y., Aaker, J., Taniguchi, A., Kazemzadeh, C., Skidmore, J. M., Martin, D. M., . . . Nakagawa, Y. (2007). Characterization of progenitor domains in the developing mouse thalamus. *The Journal of Comparative Neurology,* 505, 73-91.

Walshe, J. and Mason, I. (2003). Unique and Combinatorial Functions of Fgf3 and Fgf8 during Zebrafish Forebrain Development. *Development* 130, 4337-4349.

Wonders, C. P., & Anderson, S. A. (2006). The origin and specification of cortical interneurons. *Nature Reviews Neuroscience,* 7, 687-696.

Wullimann, M. F. and Mueller, T. (2002). Expression of Zash-1a in the Postembryonic Zebrafish Brain Allows Comparison to Mouse Mash1 Domains. *Brain Res. Gene Expr. Patterns* 1, 187-192.

Yu, M., Xi, Y., Pollack, J., Debiais-Thibaud, M., MacDonald, R. B., & Ekker, M. (2011). Activity of *dlx5a/dlx6a* regulatory elements during zebrafish GABAergic neuron development. *International Journal of Developmental Neuroscience,* 29(7), 681-691.

Zerucha, T., Stuhmer, T., Hatch, G., Park, B. K., Long, Q., Yu, G., Gambarotta, A., Schultz, J. R., Rubenstein, J. L. R. and Ekker, M. (2000). A Highly Conserved Enhancer in the *Dlx5/Dlx6* Intergenic Region is the Site of Cross-Regulatory Interactions between D*Lx g*enes in the Embryonic Forebrain. *J. Neurosci.* 20, 709-721.

Zhang, H., Hu, G., Wang, H., Sciavolino, P. $., Iler, N., Shen, M. M. and Abate-Shen, C. (1997). Heterodimerization of Msx and Dlx Homeoproteins Results in Functional Antagonism. *Mol. Cell. Biol.* 17, 2920-2932.