Identifying Dlx Regulated Genes in the Chick Craniofacial Primordia

by

Afshan Sohail

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Doctor of Philosophy
in
Molecular and Cellular Biology

Guelph, Ontario, Canada
© Afshan Sohail, September, 2022
ABSTRACT

IDENTIFYING DLX REGULATED GENES IN THE CHICK CRANIOFACIAL PRIMORDIA

Afshan Sohail
University of Guelph, 2022
Advisor: Dr. Andrew. J. Bendall

The Dlx gene family expanded via gene duplication in early vertebrate ancestors. The six paralogous genes are expressed in several vertebrate tissues, including the jaw-forming tissue of the first pharyngeal arch, where they have taken on a patterning role. Combined knockout of Dlx5 and Dlx6 in the mouse led to a homeotic transformation of the lower jaw into upper jaw structures, revealing a vital role for these homeodomain transcription factors in differentiating between upper and lower jaws. Despite the apparent importance of these genes in craniofacial development, the gene regulatory networks that they participate in are largely unknown. We describe the expression of all six Dlx genes by in situ hybridization in the pharyngeal tissue of chicken embryos between HH16 and HH26. Some candidate Dlx target genes (Pitx1, Gsc, Gbx2, Hand1 and Hand2), whose expression is lost or down-regulated in Dlx5/6 null mouse embryos, were also examined. We confirmed the endothelin dependence of each Dlx gene using Bosentan, an antagonist of the endothelin receptor. The expression of Gsc, Pitx1 and Hand2 was also reduced in md-PA1 under these conditions, consistent with the prediction that they are Dlx5 and/or Dlx6 targets. The genomic loci of these candidate Dlx-regulated genes were examined using bioinformatics tools to identify nearby sequence-conserved regulatory regions. To identify novel Dlx targets, I examined the intersection of two publicly available
datasets: p300 ChIP-seq data from E13.5 mouse craniofacial tissue (Face Base) and first arch-active enhancer regions from transgenic mouse embryo screens (VISTA). I found five enhancers from this analysis with VISTA designations mm384, mm901, mm924, mm1090, and mm1114. The orthologous chicken loci (gg) for both candidate and novel downstream genes were cloned into a GFP reporter plasmid and electroporated in ovo at HH8-10 to screen for activity in migratory CNCC or pharyngeal arch ectomesenchyme. Element gg901 showed activity in migrating neural crest at 24- and 48-hours post-electroporation (hpe), whereas gg384 was weakly active at 24 hpe. The closest genes to these enhancer regions, with known involvement in craniofacial development, are Twist1 and Bmp4, respectively. The transcription reporter assay confirmed the responsiveness of gg384, gg924, and regulatory elements associated with Gsc and Hand2, to both Dlx5 and Dlx6. In short, Dlx5 and Dlx6 appear to operate in conserved gene regulatory networks in the jaw-forming tissue of mammals and birds.
DECLARATION

I am the sole author of this manuscript and performed all the work required for this thesis project with some exceptions. A. Rebecca Walchyshyn aligned the VISTA Enhancer and FaceBase enhancer elements in the UCSC browser and retrieved the respective chicken orthologous sequence. Olivia Nicoll performed the luciferase reporter assays. Dr. Andrew Bendall performed the virus microinjections.
ACKNOWLEDGEMENTS

In the name of Allah, Most Gracious, Most Merciful.

All Praises to Allah Almighty for giving me the ability and serenity to conduct this research, the strength to persevere through hard times and for showering His countless blessings on me.

First and foremost I would like to take this opportunity to thank my advisor, Dr. Andrew Bendall. Dr. Bendall granted me the opportunity and guided me throughout this long and tedious journey with his immense support, talented advice and constructive feedback and helped me to pursue my current and future goals. I am thankful for the knowledge, depth and breadth of perspectives and diversity of ideas he brought to enrich my research enterprise. He greatly enhanced my scientific knowledge, writing and analytical skills. I cannot thank him enough in words and look forward to working with him in the future.

Secondly, I would like to extend my gratitude to my advisory committee members: Dr. Terry Van Raay, Dr. Ray Lu and Dr. Andrew Doxey for taking time out to be part of my advisory committee. I am also very grateful for their technical support, suggestion, and encouragement throughout this journey. You guys truly enriched this scientific tenure.

A big thanks to my friends, colleagues and all lab mates for providing a friendly and creative atmosphere. Special thanks to previous Bendall lab grad members Shayne and Wadood and undergrad members Hayley, Samantha, Rebecca and Olivia. I would like to extend my gratitude to other neighbouring lab members Lina, Manali, Ian, Victoria, Matt and Kholoud for bringing laughter, joy and nice coffee sessions during my stay at the campus. I met so many talented and wonderful individuals that are always ready to help.

Many people have been overwhelmingly supportive and encouraging of this work that surely my gratitude cannot be adequately expressed in just words here. Without the guidance, help, and support of those around me, completing this research would not have
been possible. I would like to acknowledge Dr. Taro Kitazawa and Dr. Masanori Uchikawa for their valuable insight and discussion of Bosentan and electroporation experiments. Also, I would like to mention the staff of the Genomic Facility namely, Jeff, Jiang and Matt. Also, I would like to mention Innes Wilson, manager of Arkell Poultry Research Facility at the University of Guelph, without his continuous support this project cannot be completed.

Finally, I would like to express my deepest gratitude to my siblings for their prayers and encouragement, even from distance. I extend my sincerest gratitude to my husband, Sohail, for his love, patience, support and encouragement which made it possible to complete this research. I must also mention my son Arham and daughters Manaal and Anaum, who bring so much joy to my life and remain a source of love, motivation and inspiration throughout my doctoral roller-coaster ride.
### TABLE OF CONTENTS

Abstract ............................................................................................................................ ii

DECLARATION...............................................................................................................iv

Acknowledgements ......................................................................................................... v

Table of Contents ........................................................................................................... vii

List of Tables .................................................................................................................. xii

List of Figures ................................................................................................................ xiii

Glossary of Abbreviations ............................................................................................. xvi

1 Introduction: Craniofacial Development, with *Dlx* Genes as a Focus .......... 1

1.1 Origin of the craniofacial skeleton ...................................................................... 1

1.2 Pharyngeal arches ............................................................................................. 3

1.2.1 The first pharyngeal arch contribution to the face primordia ..................... 4

1.2.2 Morphogenesis of arch PA2- PA6 ............................................................... 4

1.3 Signalling pathways in the first arch ............................................................... 5

1.3.1 Bone morphogenetic proteins (BMP) ....................................................... 5

1.3.2 Fibroblast growth factors (Fgfs) .............................................................. 5

1.3.3 Wnts ............................................................................................................ 6

1.3.4 Retinoic acid (RA) .................................................................................... 7

1.3.5 Sonic hedgehog (Shh) ................................................................................ 7

1.3.6 Endothelin (Edn) ........................................................................................ 8

1.4 Gene regulation during jaw patterning ......................................................... 8

1.5 *Dlx* genes ........................................................................................................ 11

1.5.1 Evolutionary history of *Dlx* genes .......................................................... 12
1.5.2  Dlx gene organization

1.5.3  Dlx gene expression and regulation by upstream signalling pathways

1.5.4  Dlx gene functions

1.5.5  Dlx gene functions in craniofacial development

1.5.6  Dlx proteins

1.5.7  Homeodomain (HD) proteins and their target DNA recognition

1.5.8  Dlx gene targets

1.6  Rational and thesis objectives

1.7  Hypothesis

1.8  Thesis Organization and Author Contribution

2  Material & Methods

2.1  Embryos and staging

2.1.1  Cell lines and transfection

2.2  Extraction of high molecular weight genomic DNA

2.3  Embryo powder

2.4  RNA extraction

2.5  Reverse transcription and preparative polymerase chain reaction (PCR)

2.6  Plasmid cloning

2.7  Riboprobe synthesis

2.8  Whole-mount In situ hybridization (WMISH)

2.9  Bosentan treatment

2.10  Quantitative polymerase chain reaction (qPCR)

2.11  Skull staining

2.12  Enhancer predictions
4.1.1 Expression patterns in PA1 ................................................................. 75
4.1.2 Effect of endothelin disruption ........................................................... 86
4.1.3 Effect of misexpressing Dlx5 ............................................................... 89
4.2 Discussion ............................................................................................... 89
4.2.1 Comparison with orthologues ............................................................ 90
4.2.2 Endothelin dependence of Dlx putative targets ................................. 91
4.2.3 Gene expression overlaps with Dlx5/6 and other Dlx paralogue pairs ... 91
4.3 Conclusion ............................................................................................... 94

5 A computational Approach to cis-Regulatory Element Discovery in the Chick: Identifying Enhancers for Cranial Neural Crest and/or the First Pharyngeal Arch 95
5.1 Results ........................................................................................................ 99
5.1.1 In silico analysis of overlapping VISTA Enhancer and p300 ChIP-Seq data 99
5.1.2 In ovo putative enhancer testing .......................................................... 101
5.1.3 Identification of Dlx regulated transcriptional enhancer upstream of Gsc 123
5.1.4 Testing the responsiveness of predicted putative enhancer elements to Dlx5 and Dlx6 ................................................................. 126
5.2 Discussion ............................................................................................... 126
5.2.1 Benchmark for putative enhancer regions ........................................... 128
5.2.2 In ovo testing ...................................................................................... 128

6 General Discussion .................................................................................... 129
6.1 Summary of work .................................................................................... 129
6.2 A proposed model of PA1 gene regulatory network in chicken ................ 131
6.3 Is Dlx5 a master regulator in PA1? .......................................................... 132
6.4 Functional redundancy vs uniqueness: Do Dlx5 and Dlx6 have different transcriptional targets? ................................................................. 133
6.5 Edn-Dlx5 related pathologies: Implication for human disease .................. 134

6.6 Future directions: ........................................................................................... 135
  6.6.1 Effect of enhancer sequence perturbation ............................................. 135
  6.6.2 Determination of genome-wide Dlx5 binding sites: .............................. 135

6.7 Significance and major contribution of research ........................................... 136

Bibliography ........................................................................................................... 138
LIST OF TABLES

Table 2.1: PCR cycling conditions ....................................................................................... 27
Table 2.2: Primers and cloning details .................................................................................. 29
Table 2.3: Proteinase K digestion conditions for chick embryos ........................................... 31
Table 2.4: qPCR primers and their respective amplicon size ............................................... 34
Table 2.5: Genomic sequences used in this study ............................................................... 38
Table 2.6: Chemicals/ reagents and their supplier’s list ...................................................... 41
Table 4.1: Selected genes and known craniofacial defects in mouse mutants ......................... 74
Table 5.1: All VISTA and Face-Base enhancers overlapping in craniofacial mouse tissue ................................................................................................................................. 103
Table 5.2: Number of GFP-positive embryos at different time windows ......................... 123
LIST OF FIGURES

Fig. 1.1: Neural crest contribution to the craniofacial structure.................................................. 3

Fig. 1.2: Schematic illustrating integrating of gene regulation in PA1 patterning cited previously in zebrafish and mouse ........................................................................................................... 11

Fig. 1.3: Possible modes of Dlx-mediated gene regulation downstream of Edn signalling in chicken PA1 ........................................................................................................................... 23

Fig. 3.1: Pharyngeal arches in chick ............................................................................................ 49

Fig. 3.2: Dlx1 general expression summary ............................................................................... 51

Fig. 3.3: Dlx1 expression in the chick embryo during pharyngeal arch patterning.................. 52

Fig. 3.4: Dlx2 general expression Summary ............................................................................... 53

Fig. 3.5: Dlx2 expression patterns in the chick embryo .............................................................. 54

Fig. 3.6: Dlx3 general expression Summary ............................................................................... 55

Fig. 3.7: Dlx3 expression patterns in the chick embryo .............................................................. 56

Fig. 3.8: Dlx4 general expression summary ............................................................................... 57

Fig. 3.9: Dlx4 expression patterns in the chick embryo .............................................................. 58

Fig. 3.10: Dlx5 general expression summary ............................................................................. 59

Fig. 3.11: Dlx5 expression patterns in the chick embryo ............................................................ 60

Fig. 3.12: Dlx6 general expression summary ............................................................................. 61

Fig. 3.13: Dlx6 expression patterns in the chick embryo ............................................................ 62

Fig. 3.14: Bosentan-treated jaw phenotype .................................................................................. 63

Fig. 3.15: Dlx gene response to reduced endothelin signalling ................................................ 64

Fig. 3.16: Differential expression of Dlx genes by qPCR analysis ............................................ 65

Fig. 3.17: Dlx family conservation among chick, mouse and dogfish shark in hinges and caps ................................................................................................................................. 68
Fig. 3.18: Schematic illustrating transformation of the lower jaw into an upper jaw.

Fig. 3.19: Schematic demonstrating possibilities of bigene cluster anomalous expression in tissue.

Fig. 3.20: A proposed model of Dlx genes regulation in md-PA.

Fig. 4.1: Gsc general expression summary.

Fig. 4.2: Gsc expression in the chick embryo.

Fig. 4.3: Pitx1 general expression summary.

Fig. 4.4: Pitx1 expression patterns of the chick embryo.

Fig. 4.5: Hand2 general expression summary.

Fig. 4.6: Hand2 expression patterns of the chick embryo.

Fig. 4.7: Hand1 general expression summary.

Fig. 4.8: Hand1 expression patterns of the chick embryo.

Fig. 4.9: Gbx2 general expression summary.

Fig. 4.10: Gbx2 expression profile.

Fig. 4.11: Differential expression of putative Dlx target gene.

Fig. 4.12: RT-qPCR following Bosentan treatment.

Fig. 4.13: Monitoring candidate gene expression following ectopic expression of Dlx5.

Fig. 4.14: Comparison of Dlx gene expression in mdPA1 with that of putative target genes.

Fig. 5.1: A general scheme of the prediction method.

Fig. 5.2: Graphical representation of pharyngeal arch-specific and non-specific enhancers in VISTA Enhancer Browser and in p300 ChIP seq. data.

Fig. 5.3: A representative overlap mapping of p300 enhancer and VISTA enhancer in UCSC genome browser.

Fig. 5.4: Electroporation of pCAGGS-EGFP control plasmid in HH9-10 chicken embryos.
Fig. 5.5: Chicken embryo demonstrating NCC migration into pharyngeal arches ....... 105
Fig. 5.6: Co-expression of reporter plasmids in cranial neural crest......................... 106
Fig. 5.7: Genomic landscaping and cloning strategy of gg901 ................................. 107
Fig. 5.8: In ovo activity of gg901 at 24 hr post-electroporation ............................... 109
Fig. 5.9: In ovo activity of gg901 at 48 hr post-electroporation ............................... 110
Fig. 5.10: In ovo activity of “antisense” gg901 at 24 hr post-electroporation ............. 111
Fig. 5.11: Genomic landscaping and cloning strategy of gg384 element .................. 112
Fig. 5.12: In ovo activity of gg384 at 24 hr post-electroporation ............................. 113
Fig. 5.13: In ovo activity of gg384 .......................................................................... 114
Fig. 5.14: Genomic landscaping and cloning strategy of gg1114 element .............. 115
Fig. 5.15: In ovo candidate enhancer gg1114 screening ......................................... 116
Fig. 5.16: Genomic landscaping and cloning strategy of gg924 element ............... 117
Fig. 5.17: In ovo candidate enhancer gg924 screening ........................................... 118
Fig. 5.18: Genomic landscaping and cloning strategy of gg1090 element .............. 119
Fig. 5.19: In ovo candidate enhancer gg1090 screening ........................................ 120
Fig. 5.20: Genomic landscaping and cloning strategy of ggHand2 element .......... 121
Fig. 5.21: In ovo candidate enhancer ggHand2 screening ..................................... 122
Fig. 5.22: Identification of a Dlx-regulated enhancer upstream of Gsc in chick genome ......................................................................................................................... 125
Fig. 5.23: Luciferase reporter assay .......................................................................... 127
Fig. 6.1: A proposed model of chick PA1 gene regulatory network ....................... 132
Fig. 6.2: Scope of current study ............................................................................... 137
GLOSSARY OF ABBREVIATIONS

ACS     Auriculocondylar Syndrome
BCA     bicinchoninic acid
BMP     Bone morphogenesis protein
ChIP-seq chromatin immunoprecipitation followed by sequencing
CNCC    cranial neural crest cells
DIG     digoxigenin
Dll     Distal-less
Dlx     vertebrate Distal-less related gene
DPE     downstream promoter element
Edn     Endothelin
Ednra   Endothelin receptor A
EDTA    ethylenediaminetetraacetic acid
eGFP    enhanced green fluorescent protein
FAS     First Arch Syndrome
FBS     fetal bovine serum
FGF     Fibroblast growth factor
FNP     frontonasal prominence
Gbx2    Gastrulation brain homeobox 2
Gsc     Goosecoid
Hand    Heart and neural crest derivatives expressed transcript
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>HD</td>
<td>homeodomain</td>
</tr>
<tr>
<td>HDRE</td>
<td>homeodomain response element</td>
</tr>
<tr>
<td>HEK 293T</td>
<td>human embryonic kidney 293T</td>
</tr>
<tr>
<td>HH</td>
<td>Hamburger and Hamilton stage</td>
</tr>
<tr>
<td>IDP</td>
<td>intrinsically disordered protein</td>
</tr>
<tr>
<td>IDT</td>
<td>Integrated DNA Technology</td>
</tr>
<tr>
<td>KO</td>
<td>knockout</td>
</tr>
<tr>
<td>KOH</td>
<td>potassium hydroxide</td>
</tr>
<tr>
<td>LB</td>
<td>Luria Bertani</td>
</tr>
<tr>
<td>LiCl</td>
<td>lithium chloride</td>
</tr>
<tr>
<td>MC</td>
<td>Meckel's cartilage</td>
</tr>
<tr>
<td>md</td>
<td>mandibule</td>
</tr>
<tr>
<td>md-PA1</td>
<td>mandibular branch of the first pharyngeal arch</td>
</tr>
<tr>
<td>Mef2</td>
<td>Myocyte enhancer factor 2</td>
</tr>
<tr>
<td>MRGs</td>
<td>master regulatory genes</td>
</tr>
<tr>
<td>mx</td>
<td>maxillar</td>
</tr>
<tr>
<td>mx-PA1</td>
<td>maxillary branch of the first pharyngeal arch</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Center for Biotechnology Information</td>
</tr>
<tr>
<td>NCC</td>
<td>neural crest cells</td>
</tr>
<tr>
<td>PA</td>
<td>pharyngeal arches</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffer saline</td>
</tr>
</tbody>
</table>
PCR  polymerase chain reaction  
PEI  polyethylenimine  
PFA  paraformaldehyde  
PIC  pre-initiation complex  
Pitx1  Paired-like homeodomain transcription factor 1  
PQ  palatoquadrate  
RA  retinoic acid  
Res  responsive elements  
SDH  succinate dehydrogenase  
SDS  sodium dodecyl sulphate  
SEC  super elongation complexes  
Shh  Sonic hedgehog  
SPF  specific pathogen-free  
TAD  topologically-associated domains  
TF  transcription factor  
tk  thymidine kinase  
Wnt  Wingless-related integration site
Chapter 1

Introduction: Craniofacial Development, with Dlx Genes as a Focus

The modern vertebrate head is a prerequisite for transitioning from filter-feeding to active and more effective predation (Northcutt, 2005). This diversification from (jawless) agnathans to (jawed) gnathostomes was driven by developmental innovations, including a combination of orchestrated and discrete genetic events, such as the diversification of neural crest cells, ectodermal placodes, new regulatory pathways and articulated jaws (Depew et al., 2002; Kraus & Lufkin, 2006).

1.1 Origin of the craniofacial skeleton

Neural crest cells (NCCs) are a significant modern vertebrate innovation (Hunt et al., 1991). They originated from the edge of the neural plate and are found at the border of the neural fold (Clouthier et al., 2010). NCCs can be characterized into five groups that form different lineages. Cranial NC cells (CNCCs) form glia and connective tissues, facial cartilage, teeth, middle ear and jawbones. Cardiac NC cells form the aorta and pulmonary artery, cartilage and connective tissue of pharyngeal arches 3, 4, and 6. Trunk NC cells form melanocytes, ganglia and sensory neurons. Lastly, vagal and sacral NC cells form the parasympathetic ganglia of the gut (Betancur et al., 2010). NCCs delaminate and migrate to different tissues as the neural plate grows and elongates. Cranial NCCs form in the dorsal mid- and hindbrain and then migrate to different destinations within the cranial region. When these cells migrate, topographic information is limited, and patterning signals that emanate from the ectoderm and pharyngeal endoderm act on
these CNCCs (Chai & Maxson, 2006; Jeong et al., 2008; Santagati & Rijli, 2003). For instance, tissue-specific endothelin 1 signalling acts during the migration and differentiation of NCCs to confer a regional maxilla-mandibular identity (Couly et al., 2002; Noden & Trainor, 2005; Ruhin et al., 2003). The transient NCCs that have emigrated from the mid-and hindbrain then differentiate to generate bone, and connective tissues (Cesario et al., 2015; Clouthier et al., 2010; Couly et al., 1998; Köntges & Lumsden, 1996; Noden, 1988; Noden & Trainor, 2005; Noden, 1983; Schilling & Kimmel, 1994).

Craniofacial development is a three-staged process defined by the (1) formation and induction of CNCCs (2) migration and entry of CNCCs into the pharyngeal arches (PA) and (3) NCC interaction and interpretation of signalling cues (Cesario et al., 2015).

Early in development, the mammalian skull is a cartilaginous structure termed the chondrocranium (Hanken & Thorogood, 1993). This initial structure gives rise to three distinct derivatives, the neurocranium, dermatocranium, and splanchnocranium (Fig. 1.1) (or viscerocranium) (Hanken & Thorogood, 1993). The neurocranium facilitates the development of the structures around the nose, inner ear and skull base. The dermatocranium appears during intramembranous ossification and covers the other two units of the chondrocranium (Hanken & Thorogood, 1993; Kraus & Lufkin, 2006). The splanchnocranium gives rise to pharyngeal arches and facial structures. The mammalian splanchnocranium has two major parts that help define the jaws: the derivative of the palatoquadrate (PQ, mx-PA1 derivative) and Meckel's cartilage (MC, md-PA1 derivative). Neural crest-derived structures are shown in Fig. 1.1.

All derivatives of the splanchnocranium undergo ossification, except for Meckel's cartilage. The splanchnocranium originates from CNCCs, originally derived from the anterior hindbrain and midbrain regions (Beverdam et al., 2002). It is believed that signals from the endoderm and ectoderm provide positional information to the CNCCs. CNCCs originate from two distinct areas in the skull: the rostral and caudal Hox-negative region. The CNCCs from rhombomere 2 populate the PA1 whereas CNCCs from rhombomere 4 populate the PA2 (Beverdam et al., 2002). CNCCs migrate from the Hox-negative regions
of the dorsal brain to form ectomesenchyme. Ectomesenchyme contributes to the patterning of PA whereas frontonasal prominence (FNP) forms the middle and upper facial skeleton. The proximal half of PA1 develops into the maxillary arch (mx-PA1), and the distal half develops into the mandibular arch (md-PA1). The second arch, PA2, gives rise to the structures of the inner ear and neck skeleton (Köntges & Lumsden, 1996; Solomon & Fritz, 2002).

![Fig. 1.1: Neural crest contribution to the craniofacial structure.](image)

(A) Chick embryo illustrating neural crest mesenchyme derived structures in brown and mesodermal mesenchyme region is shown in grey. Adapted and modified from (Kuratani, 2005) (B) Bird jaw skeleton at stage 40. The blue color represents the NC-derived structures, orange shows the mesodermal mesenchyme structures. Modified from (Woronowicz & Schneider, 2019). The stippled portion black depicts the neurocranium, red stippled region is derived from splanchnocranium, and the green stippled part points to the dermatocranium. di, diencephalon; fnp, frontonasal prominence; ey, eye; md, mandible (lower jaw); mes, mesencephalon; mx, maxilla (upper jaw); pa, palatine; pharyngeal arches (I, II, III, IV); pm, pre-maxilla; qd, quadrate; sp, splenial; sq, squamosal; su, surangular; tel, telencephalon.

### 1.2 Pharyngeal arches

The PAs are transient conserved evolutionary structures in the vertebrate embryo that develop to form various primordia in the face and neck. The arches are associated with three unique developmental events involving neural crest cells: migration from
rhombomeric neuroectoderm to the pharyngeal arches, proliferation as an ectomesenchyme, and differentiation into terminal structures (Clouthier et al., 2000). There are 5 pairs of PAs in mammals, numbered 1-4 and 6. The PAs comprise a mesenchymal core of cranial mesoderm, lined with surface ectoderm lining the pharyngeal cleft or groove and pharyngeal endoderm facing the pharynx and oral cavity. Each arch is provided with a PA artery and cranial nerve deeply rooted in the mesenchyme. NC-derived mesenchyme is sandwiched between the ectoderm and endoderm of the PA and is referred to as ectomesenchyme. As development progresses, NC-derived mesenchyme migrates and populates the PA, giving rise to major craniofacial structures as well as face and neck connective tissue. The NCCs from the caudal midbrain and rostral hindbrain migrate to PA1.

1.2.1 The first pharyngeal arch contribution to the face primordia

The first pharyngeal arch is the site of jaw patterning. PA1 is divided into the maxillary (mx) branch and mandibular (md) branch. Maxillary PA1 contributes to all bones of the upper jaw, including the maxilla and palatine, and the incus bone of the inner ear (Chai et al., 2000; Noden, 1978). Mandibular PA1 forms the lower jaw, including Meckel's cartilage, the intramembranous dentary bone, and the malleus, another bone of the inner ear.

1.2.2 Morphogenesis of arch PA2- PA6

The second arch forms neck and middle ear bones (stapes) and facial muscles (Köntges & Lumsden, 1996). PA3 forms the hyoid bone, stylopharyngeus muscles, tongue, pharynx, inner ear bone, and incus (Cesario et al., 2015). PA4 and PA6 contribute to forming the larynx, thyroid, arytenoid, cricoid cartilage and pulmonary artery (Noden & Trainor, 2005). These arches also contribute to the formation of laryngeal and pharyngeal muscles. In mammals, it is believed that PA5 does not form or degenerate soon after it appears (Cesario et al., 2015).
1.3 Signalling pathways in the first arch

Several signalling pathways are implicated in craniofacial morphogenesis, such as those regulated by Bone Morphogenesis Protein (BMP), Wnt (wingless-related integration site), Fibroblast Growth Factor (FGF), Retinoic acid, sonic hedgehog (Shh) and Endothelin (Edn) ligands.

1.3.1 Bone morphogenetic proteins (BMP)

BMPs are members of the transforming growth factors β superfamily of secreted ligands. Other secreted members include nodal, activins, and growth & differentiation factors. These ligands bind to serine/threonine receptors, which subsequently activate SMAD proteins and, as a result, regulate transcriptional targets (Massagué, 2012). BMP gradients and their antagonism by noggin, chordin and follistatin are required for epidermal fate and neural crest induction (Betancur et al., 2010, 2011). BMP is primarily involved in maxillary and FNP morphogenesis (Lee et al., 2001) but can also contribute to the growth of mandibular prominence (Mina et al., 2002; Richman & Lee, 2003; Santagati & Rijli, 2003). In the chick embryo, BMP is expressed in both branches of PA1 in a time-dependent manner around HH20-22. In the early mandibular developmental stages of the chicken, the fine gradient and spatial control of BMP and FGF are crucial for mesenchymal survival and death. BMP7 promotes apoptosis in the lateral region of the mandibular arch with ectopic expression of Msx1 (Mina et al., 2002). BMP signalling also shapes the upper palatal shelves (Santagati & Rijli, 2003). Blocking BMP expression with Noggin results in transforming the maxillary process into FNP. BMP7 also plays a vital role in epithelial seam degradation, which leads to lip fusion (Ashique et al., 2002).

1.3.2 Fibroblast growth factors (Fgfs)

Fgfs are implicated in almost every stage of development ranging from implantation to organogenesis in mammals and birds. FGF receptors are tyrosine kinases and activate a cascade reaction by activating Ras, Raf, ERK and MEK 1/2 pathways (Dorey & Amaya,
Like the BMP signalling pathway, Fgfs are also involved in NCC induction. The Fgfs, together with BMP signalling, regulate the expression of Snail2, an early marker of NCC. Moreover, Fgfs also control Msx gene expression (Mayor et al., 1997). The Fgf and FgfR signalling pathways have functional redundancy in NC development (Jones & Trainor, 2005). In chickens, Fgf is required to develop the FNP and maxillary processes (Szabo-Rogers et al., 2008). Furthermore, disruption in Fgf signalling in chickens around Hamilton Hamburger stage 17 (HH17) results in a shortening of the upper beak (Hu & Marcucio, 2009). Fgf2 and Fgf4 convey signals from the epithelium to the medial region to promote the growth of mandibular processes and Meckel’s cartilage (Mina et al., 2002). High and low Fgf concentration along the anterior-posterior axis is responsible for Lhx and Gsc transcription factor (TF) regulation, respectively (Tucker et al., 1999). Interestingly, Gsc is also downstream of the endothelin signalling pathway. Fgf and BMP act antagonistically, both spatially and temporally. In mice, a mutation in FgfR1 prevents NCCs migration into PA2 (Trokovic et al., 2003).

1.3.3 Wnts

Wnts, vertebrate homologues of wingless in Drosophila were initially discovered as a mammalian proto-oncogene. The WNT ligand binds to several different receptors through canonical and non-canonical pathways. In canonical pathways, the WNT ligand binds to a Frizzled receptor and triggers the translocation of β-catenin into the nucleus, further regulating downstream transcriptional targets. The canonical and non-canonical WNT signalling pathways both play a crucial role in craniofacial development by protecting the NCC from apoptosis (Maretto et al., 2003). Wnt5a, a non-canonical ligand, is expressed in PA1 and PA2, which are prominent regions that contribute to the formation of craniofacial structures. Mutation or interruption in the WNT pathway, more specifically in non-canonical pathways, results in lip and palate clefting. For normal craniofacial cartilage differentiation, activation and repression of non-canonical and canonical pathways are required (Brugmann et al., 2007; Fuchs et al., 2009; Geetha-Loganathan et al., 2009; Phillips et al., 2012). The non-canonical pathway includes activating calcium-dependent or small G-proteins, RhoA and Rac1. Wnt5a expression is high in the maxilla, mandible,
and FNP and directs Meckel's cartilage placement and length. Wnt11, another non-canonical pathway paralogue, controls the proliferation of maxillary cells and thereby regulates maxillary size. In the mesenchyme, Msx1, Msx2, and Lhx8 are direct transcriptional targets of the canonical Wnt pathway (Malt et al., 2014; Song et al., 2009).

1.3.4 Retinoic acid (RA)

RA is a derivative of vitamin A and is critically involved in forebrain and hindbrain development (Rhinn & Dollé, 2012). In mice, RA is synthesized in pharyngeal arch 3 to pharyngeal arch 6 and is regarded as a posteriorizing signal (Durston et al., 1989; Rhinn & Dollé, 2012). PA 3, 4, and 6 and their derived structures fail to develop in RA pathway mutants. Abe and colleagues (2008) have illustrated that RA indirectly controls the identity of PA1. Administration of RA into pregnant mice at stage E8.5, when NCCs are migrating, induced a homeotic transformation of the mandibular arch into the maxillary arch with the subsequent inhibition of Fgf signalling, more specifically Fgf8 in arch ectoderm (Abe et al., 2008). This shows that RA controls the Fgf gradient in migrating NCCs. Additional genes that showed downregulation within the mandibular arch were Dlx5 and Dlx6, Pitx1, and Hand2, whereas expression of Alx4 remained the same. The phenotype of RA-treated embryos differed slightly in bone structure between treatment groups than those of the homozygous Dlx5/6−/−, Edn1−/− and Ednra−/− mutant embryos (Abe et al., 2008). In the chick, RA-treated embryos showed defects in limb and orofacial morphogenesis with duplication of digits and completely truncated upper beak and nostrils (Tamarin et al., 1984).

1.3.5 Sonic hedgehog (Shh)

Hedgehog (Hh), like WNT, was first discovered in Drosophila as being involved in body segmentation (Nüsslein-volhard & Wieschaus, 1980). In mammals, three homologues of Hh, Sonic hedgehog (Shh), Desert hedgehog (Dhh) and Indian hedgehog (Ihh), have been discovered (Echelard et al., 1993). Shh is the only vertebrate hedgehog expressed within the pharyngeal arch and impacts the vertebrate brain, tissue patterning, limb morphogenesis, lungs and teeth (Jeong et al., 2004; Moore-Scott & Manley, 2005). Shh,
expressed within the epithelium of the PA, contributes to the development of the FNP and maxillary structures but not to the mandibular prominence (Marcucio et al., 2005). Inhibition of Shh in PA1 leads to loss of expression of several Fox genes (Jeong et al., 2004). In the absence of Shh, the FNP fails to develop, whereas the introduction of Shh-secreting fibroblast cells in the FNP results in an ectopic upper beak in the chick (Santagati & Rijli, 2003). Inhibition or ablation of the Shh signalling pathway in NCC resulted in increased apoptosis and decreased proliferation (Jeong et al., 2004).

1.3.6 Endothelin (Edn)

Endothelin proteins (Edn) were first discovered as an endothelium-derived factor which activates vasoconstriction by narrowing the blood vessels (Yanagisawa et al., 1988). These proteins bind to the G-protein receptors, G\textsubscript{αq} and G\textsubscript{α11}, and activate the calcium and diacylglycerol (DAG) signalling pathway (Clouthier et al., 2010; Horinouchi et al., 2013).

In vertebrates, endothelin is extensively involved in PA1 development and is secreted by the epithelium and mesodermal core of PA. Endothelin confers the maxillary and mandibular identity to the PA1 and genetically programs the lower jaw. In mice, endothelin downstream transcriptional targets include Gsc, Dlx5, Dlx6, Dlx3, Hand1, Hand2 and Pitx1 (Clouthier et al., 1998; Ozeki et al., 2004; Ruest et al., 2004; Thomas et al., 1998). Among them, Dlx5 and Dlx6 are early effectors that regulate other TF in PA1. Endothelin signalling will be discussed in greater detail in chapter 3.

1.4 Gene regulation during jaw patterning

A number of insights have been gained from work with mice and zebrafish that sheds some light on the gene regulatory networks that underpin jaw patterning. Depew et al (2002) highlight the differentially expressed genes in wild-type and Dlx5/6 knockout mice and shows the interplay between signalling pathways. We know Dlx1, Dlx2, Dlx5, and Dlx6 contribute to upper and lower jaw morphogenesis. Depew and colleagues hypothesized that the signal for patterning emanates from the ectoderm (Chai & Maxson,
2006; Depew et al., 2002; Santagati & Rijli, 2003), where Fgf8 is expressed at the border of the maxillary and mandibular arches. Dlx1 and Dlx2 in mx-PA1 and Dlx5 and Dlx6 in md-PA1 are expressed as a result of this signalling. In the absence of Dlx5/6, only Dlx1/2 interacts with the Fgf8 signal, resulting in the subsequent duplication of the jaw. Furthermore, Depew and colleagues hypothesized that, in the mandibular arch, Dlx5 and/or Dlx6 regulate Pitx1, Dlx3, Alx4, Barx1 and Hand2. Whereas, outside of the mandibular domain, those same TFs inhibit the expression of Prx1 and Prx2 (Depew et al., 2002).

Vanyai et al (2019) used RNA-seq and ChIP-seq analysis to focus on maxillary and mandible programming by Dlx genes. This model proposes that Dlx1 and Dlx2 regulate the expression of Pou3f3 in the maxillary arch for its patterning, whereas, Dlx5 and Dlx6 regulate a plethora of TFs, such as Dlx3, Dlx4, Barx1, Hand1, Hand2, Gbx2, Gsc, Osrt1/2, sim2, Dlx1as, and Dlx6. Another unique insight of Vanyai et al. (2019) is the indirect regulation of Dlx5 and Dlx6 in the mandibular arch by Mef2C.

According to Clouthier et al (Clouthier et al., 2010), mandibular arch patterning revolves around Ednra signalling in mice and zebrafish. In zebrafish, there are two Ednra receptors, Ednra1 and Ednra2. An interaction between Ednra1 and Edn1 initiates a series of reactions resulting in the expression of dlx5 and dlx6. Ednra and Mef2C regulate Dlx5 expression in both mice and zebrafish. In the Clouthier model, dlx5 and/or dlx6 regulates hand2 and dlx3 directly, and hand2 further restricts the expression of gsc. Furthermore, Ednra directly controls Hand1 expression in mice and msxE and ephA2 in zebrafish. Clouthier and colleagues divide the mandibular arch into ventral/distal, intermediate, and dorsal/ proximal domains. Ednra inhibits the expression of Bapx1 in the intermediate domain of md-PA1 and mediates inhibition of twist1, dlx1, dlx2, wnt5a, and eng2 genes in the dorsal domain (Clouthier et al., 2010).

Jeong et al. (2008) defined three groups of genes that are expressed within the craniofacial primordia. Group A comprises the mandibular jaw-specific genes and includes Dlx3, Dlx4, Hand1, Hand2, Gbx2, Gsc, Alx3, Alx4, Bmpcr, Cited1, Zac1, Unc5c,
Hgf, Rgs5, A/S Dlx1 and Evf1/2. Group B comprises the maxillary jaw-specific genes such as Pou3f3, 2610016I09Rik and 2900092D14Rik. Lastly, Group C includes genes that are expressed at the junction of md- and mx-PA1 or in the zygomatic arch but have more of a maxillary component than mandibular. Some genes of group C are Foxl2, E330015D05Rik, and Irx5. Group C genes are also expressed in other cranial derivatives, such as sensory organs. Although Group C genes were shown to be differentially expressed upon loss of both Dlx5 and Dlx6, most have no known jaw defects when knocked out themselves.

Jeong et al. (2008) proposed that, in md-PA1, Dlx5 and Dlx6 promote the expression of Group A TFs, while simultaneously repressing the expression of group B and C genes. Dlx1 and Dlx2 expression in the mandibular arch is partially redundant and also regulates the expression of Group A genes. Both Dlx1 and Dlx2 promote and regulate the expression of group B and C genes localized in the maxillary arch, the border of the maxillary and mandible arches, and the sensory organs. Being based on DNA microarray data, this study did not provide any information about direct or indirect targets of Dlx5 and Dlx6 in md-PA1 or of regulation of Dlx genes during craniofacial morphogenesis (Jeong et al., 2008). The data described herein show both conserved gene regulatory interactions, as well as species-specific differences (Fig. 1.2). All studies show a central role for Dlx5 and Dlx6 in md-PA1. Depew et al. (2002) proposed that a maxillary identity represents the default, upon which lower-jaw fate is superimposed. As far as jaw morphogenesis is concerned, the endothelin pathway contributes a primary role to arch identity by supporting Dlx5 and Dlx6 gene expression during a critical window of jaw patterning. These regulatory TFs activate and repress target genes and endow the mandibular arch with its unique identity. Although a few players are known, we do not have complete knowledge about the jaw gene network, which adds to the complexity and challenge of understanding jaw-specific gene regulatory mechanisms.
Pathways or genes that were conserved among mice and zebrafish are shown in purple. Genes that are only implicated in zebrafish are shown in blue, whereas genes only involved in mice are shown in orange color. This figure is not to scale.

1.5 Dlx genes

Dlx genes are homologues of the Distal-less (Dll) gene and are a subset of homeobox genes, with a common homeobox sequence of 180 bp (Holland & Takahashi, 2005), are an early-acting group of developmental regulatory genes (Merlo et al., 2000) and were first discovered in the fruit fly Drosophila (Depew et al., 2002, 2005; Ferguson et al., 2014; Holland & Takahashi, 2005). Homeobox genes are conserved in almost all multicellular animals and plants (Holland, 2013; Mukherjee et al., 2009). In Drosophila, the single Dll
gene is involved in antenna and appendage development. \emph{Dll} genes have an ancient origin near the base of the metazoan tree (Stock et al., 1996) and emerged as critical embryonic genes and play crucial roles in all pivotal stages of development, such as proliferation, differentiation, migration and cell adhesion (Cobos et al., 2007; Colasante et al., 2008; Lee et al., 2003; Lee et al., 2005; McKeown et al., 2005). Evolutionarily, the \emph{Dlx} gene family profoundly impacted the embryonic development of the vertebrate head (Zerucha et al., 2000).

1.5.1 Evolutionary history of \emph{Dlx} genes

Invertebrates, including basal chordates, usually have a single \emph{Dll} gene (some tunicates exhibit two), whereas mammals have six \emph{Dlx} genes (Stock et al., 1996). It is proposed that the duplication of an ancestral \emph{Dlx} gene, followed by two rounds of genomic duplication, could have accounted for the present mammalian \emph{Dlx} gene family (Ellies et al., 1997). The duplication theory is also consistent with the overlapping expression and functional redundancy among the members of this gene family (Ellies et al., 1997). Jawed and jawless vertebrates had a common ancestor 500 million years ago. Genomic analysis of conserved elements provides a clue that Edn signalling was in place before the evolutionary "genomic duplication" event occurred (Square et al., 2020).

1.5.2 \emph{Dlx} gene organization

In vertebrates, six paralogous genes have been identified, located in bigene clusters of \emph{Dlx1} and \emph{Dlx2}, \emph{Dlx3} and \emph{Dlx4}, and \emph{Dlx5} and \emph{Dlx6}, respectively (Debiais-Thibaud et al., 2013; Stock et al., 1996). Paired \emph{Dlx} genes are organized in an inverted (tail-to-tail) configuration and are separated by a short, 3.5-16 kb intergenic region. \emph{Dlx} genes can be further divided into two subfamilies based on sequence homology. One subfamily comprises \emph{Dlx1}, \emph{Dlx4}, and \emph{Dlx6}, and the other subfamily comprises \emph{Dlx2}, \emph{Dlx3}, and \emph{Dlx5} paralogues (Takechi et al., 2013). Thus, each member of a linked \emph{Dlx} pair shows more sequence homology to genes in other linked pairs rather than the \emph{cis}-linked parologue (Stock et al., 1996). Among all \emph{Dlx} paralogues, \emph{Dlx1} shows the highest level of similarity with the Drosophila \emph{Dll} gene (Holland & Takahashi, 2005; Stock et al., 1996). \emph{Dlx} genes
consist of three exons and two introns (Ellies et al., 1997), a typical exon-intron organization shared by all chordates (Stock et al., 1996). The homeobox is interrupted by intron 2 (Ellies et al., 1997). All coding domains of genes show more sequence conservation than non-coding regions (Sumiyama & Ruddle, 2003).

1.5.3 *Dlx* gene expression and regulation by upstream signalling pathways

*Dlx* genes are responsive to a variety of upstream signalling pathways, depending on the cellular context. Shh can induce *Dlx2* expression in the forebrain (Debiais-Thibaud et al., 2013; Panganiban & Rubenstein, 2002). In addition, BMP2 can induce *Dlx2* expression in chondrocytes (Kraus & Lufkin, 2006). BMP4 triggers *Dlx5* expression in osteoblasts, *Dlx1* and *Dlx2* expression in dental mesenchyme, and *Dlx3* expression in embryonic ectoderm (Kraus & Lufkin, 2006). FGFs are also known to induce and sustain the expression of all *Dlx* gene family members. Specifically, FGF2 can induce *Dlx3* and *Dlx5* expression in the chicken limb (Fallon et al., 1994). FGF19 with Wnt8c can induce the expression of *Dlx5* in the inner ear (Kraus & Lufkin, 2006). In mice, FGF8 induces *Dlx1* and *Dlx2* expression in dental mesenchyme and *Dlx1* expression in the chicken mandibular and hyoid pharyngeal arch (Shigetani et al., 2002). Edn1 protein is known to induce the expression of some *Dlx* genes, especially *Dlx5* and *Dlx6* in md-PA1 (Medeiros & Crump, 2012).

1.5.3.1 *Dlx5* regulation by monocytic leukemia zinc-finger (MOZ)

Vanyai et al. (2019) showed that one of the histone lysine acetyltransferase genes, monocytic leukemia zinc-finger (Moz), directly affects developmental processes such as cardiogenesis and craniofacial morphogenesis. In Moz knockout mice, they observed shortening of the mandible and its associated bone structures. Vanyai and colleagues also showed the direct occupancy of Moz protein on the *Dlx5* locus by ChIP-sequencing analysis (Vanyai et al., 2019). They further found that some downstream genes such as *Dlx1, Dlx2, Dlx3, Gbx2, Sim2* and *Osr2* showed differential expression in Moz knockout mice, but the genes upstream of Moz gene did not show any altered expression.
Moreover, they concluded that Moz has to be expressed in the NCC to directly regulate Dlx5 (Vanyai et al., 2019).

Moz is required for the acetylation of lysine, specifically H3K4Ac, on the Dlx5 promoter (Vanyai et al., 2019). Mutation in the MOZ gene halts the acetylation of H3K4. Though the exact role of H3K4Ac is unknown, it is known to be associated with active gene transcription in two ways. First, acetylation neutralizes the charge on lysine residues, which results in reduced DNA-histone interaction. This phenomenon facilitates the binding of other proteins involved in transcription. Secondly, H3K4Ac is involved in the recruitment of super elongation complexes (SEC) (Gates et al., 2017; X. Liu et al., 2015). These SEC facilitate transcription by promoting the transition from initiation to the elongation phase (Gates et al., 2017). The inability of SECs to do so will result in truncated mRNA, consequently halting the transcription process and regular gene expression (Gates et al., 2017).

1.5.3.2 Myocyte enhancer factor 2C (Mef2C)-mediated regulation of Dlx5 and Dlx6 expression

The myocyte enhancer factor 2 (Mef2) takes part in craniofacial development and melanocyte differentiation (Agarwal et al., 2011; Verzi et al., 2007). In vertebrates, four homologues, Mef2A, Mef2B, Mef2C, and Mef2D, are expressed (Verzi et al., 2007). These proteins homodimerize or heterodimerize to bind DNA at a specific site called a Mef2-site to regulate the expression of target genes. Among all homologues, Mef2C is the only member of the Mef2 family expressed in the NC and is involved in craniofacial development. Inactivation of Mef2C resulted in upper airway obstruction, delayed ossification, and hypoplasia of the mandibles (Verzi et al., 2007).

Moreover, in Mef2C knockout (KO) mice, several skull and craniofacial bones and the tympanic ring were absent. Interestingly, Dlx5/6 and Hand2 expression were also lost in conditional Mef2C KO mice, which shows that Mef2C is an upstream regulator. In mice, Mef2C is expressed in the NC of PA around E9.5. Furthermore, Verzi and colleagues reported a deeply conserved enhancer element upstream of the Dlx6 coding region (Verzi
et al., 2007). This region is responsible for the expression of the Dlx5-Dlx6 gene pair. This non-coding enhancer element also contains multiple Dlx-binding sites. The presence of Dlx-binding sites suggests autoregulation by Dlx5 and Dlx6. Indeed, Mef2C acts with Dlx proteins to synergistically activate transcriptional targets (Verzi et al., 2007). In osteoblasts, Dlx5 and Mef2C synergistically regulate Runx2 expression, directly binding to a novel enhancer in an enhanceosome (Kawane et al., 2014).

1.5.4 Dlx gene functions

All members of the Dlx gene family perform diverse functions. Dlx genes also play a role in chicken feather morphogenesis and, at the same time, positively and negatively regulate the feather formation by Fgf and BMP signalling pathways, respectively. Dlx2, Dlx3 and Dlx5 genes are expressed in the dermis of chicken embryos and regulate tenascin and NCAM molecules, which are essential for feather bud initiation and growth (Rouzankina et al., 2004). Also, Dlx2 directly regulates genes responsible for tangential migration of GABAergic interneurons (Anderson et al., 1997); Dlx3 is involved in tissue mineralization (Samee et al., 2007). Dlx1 and Dlx2 are involved in olfactory bulb development, whereas Dlx5 and Dlx6 are expressed in the otic vesicles (Sajan et al., 2011) and also regulate chondrogenesis (Hsu et al., 2006; Robledo et al., 2002; Zhu & Bendall, 2009). Esx1, a paired-like homeodomain protein that plays a pivotal role in placental tissue morphogenesis, is regulated by Dlx3 and is responsible for embryonic survival (Morasso et al., 1999; Roberson et al., 2001).

1.5.5 Dlx gene functions in craniofacial development

Dlx genes have an extensive impact on craniofacial development. This gene family defines the proximal-distal identity of the PA1. Dlx1 and Dlx2 are expressed in the proximal or upper jaw, whereas all Dlx genes, Dlx1-Dlx6, are expressed in the lower jaw (Beverdam et al., 2002). Dlx3, Dlx5, and Dlx6 are Edn1-dependent in the lower jaw in mice (Ozeki et al., 2004). Dlx1 and Dlx2 in the upper jaw (maxilla) are involved in upper jaw morphogenesis, whereas in the lower jaw (mandible), they are engaged in lower arch
patterning, which shows $Dlx$ genes have a context-dependent role in the tissue in which they are expressed (Jeong et al., 2008).

Both gain-of-function and loss-of-function strategies have been used to define the $Dlx$ code in the PA1 and the genetic pathway(s) downstream of Dlx5 and Dlx6 in md-PA1. The md-PA-derived structure, MC, is very short in Dlx5 mutant mice; malformation starts around E13.5 and can be observed clearly at E15.5. The mutant MC grows proximal to the dentary, where it splits into two branches and forms an ectopic strut towards the pterygoid and basisphenoid (PA1 bones); the other branch extends into the malleus. The temporalis, a lower jaw bone, is attached to the coronoid process and runs parallel to the mandibular PA1. In Dlx5 mutant mice, the molar alveoli are shorter, more constricted distally and broader proximally, while the condylar and angular processes are short and juxtaposed (Depew et al., 2005). Upper palate clefting and defects in mx-PA1 are also observed in $Dlx5$ mutants, and the pterygoids show abnormal chondroid bone and secondary cartilage. $Dlx5$ is also expressed in dental mesenchyme, and both mandibular and maxillary teeth are malformed in $Dlx5$ mutants; their crowns are poorly mineralized, and the incisors are short or distorted (Depew et al., 1999). $Dlx6^{-/-}$ mutants are born alive but die shortly after birth due to excessive air inhalation (Jeong et al., 2008). $Dlx6^{-/-}$ mutants also demonstrate short mandibles, dentary bones lacking the coronoid processes, and gonial bones attached to an ectopic piece of bone named the os-paradoxicum (Depew et al., 1999). The homozygous double $Dlx5/6^{-/-}$ mutants showed severe defects in the lower jaw relative to single mutants. $Dlx5/6^{-/-}$ mutants failed to develop MC, mandible and calvaria structures. Loss of two alleles (either $Dlx5^{+/-}; Dlx6^{+/-}$ or $Dlx5^{-/-}; Dlx6^{+/-}$) resulted in the loss of the coronoid process and the shortening of mandibles and loss of the angular processes. Removal of three alleles ($Dlx5^{-/-}/Dlx6^{+/-}$) directs the loss of bones such as; coronoid, condylar and angular processes and severe shortening of the mandibles. Finally, the maxillary structure was identifiable but significantly distorted in double KO mutants ($Dlx5^{-/-}/Dlx6^{+/-}$). In contrast, the mandibular structure of the jaw becomes indistinguishable from the maxillary arch or upper jaw. This transformation of the lower jaw into the upper jaw suggests a unique and redundant
function of $Dlx5$ and $Dlx6$ (Depew et al., 1999). However, it is noteworthy that, while severe defects can be seen in $Dlx1/2$ knockouts, the maxillary jaw does not lose its identity, whereas, in $Dlx5/6$ double knockout mice, the mandibular jaw undergoes a homeotic transformation.

Conversely, in a $Dlx5/6$ overexpression model, maxillary identity is transformed into a mandibular fate by altering the proteomic signature. Shimizu and colleagues (2018) showed that downregulated genes in the mdPA1 of $Dlx5/6^{-/-}$ mutants were upregulated in the transformed maxillary arch of mice with $Dlx5/6$-overexpression in NCCs. Presumably, these sets of genes were direct or indirect targets of Dlx5/6 in the mandibular arch. Furthermore, these genes were additionally identified as mandibular jaw markers (Shimizu et al., 2018). In the absence of one paralogue, the other paralogue recapitulates the function/phenotype, highlighting their functional redundancy and their allelic equivalency (Bendall, 2016; Vieux-Rochas et al., 2010; Zhu & Bendall, 2009).

Depew et al. (2002) discovered that functional redundancy in $Dlx1$ and $Dlx2$ mutants limits defects in the upper jaw while insignificant changes were observed in the lower jaw. $Dlx1$ null mice showed defects in mx-PA1-derived bones, including palatine, lamina, jugal, squamosal, alisphenoid and pterygoid (Depew et al., 2002). In $Dlx2$ mutants, maxillary bones (palatine and medial bone) were reduced in size. The double mutant $Dlx1/2^{-/-}$ mice not only showed the single-gene KO phenotype but also exhibited new defects, including the absence of upper molar teeth. Notably, no distal defect is observed in $Dlx1/2^{-/-}$ mice, which suggests the functional redundancy of the $Dlx$ genes in the distal or mandibular portion where $Dlx5$ and $Dlx6$ are co-expressed (Merlo et al., 2000).

Further evidence of functional equivalency came from Gordon et al. (2010). In the chick embryo, $Dlx2$ is involved in forming the jugal bone by inducing undifferentiated mesenchymal aggregation. Ectopic expression of $Dlx2$ and $Dlx5$ showed the same phenotype in the upper jaw of an avian embryo, highlighting the conservation of function of $Dlx$ genes in birds as well (Gordon et al., 2010).
1.5.6 Dlx proteins

1.5.6.1 Protein structure

All Dlx proteins share a nearly identical DNA-binding domain, the homeodomain, consisting of 60 amino acids comprising three α-helices (Piper et al., 1999). The third helix also called the "recognition helix," confers sequence-specific recognition in the major groove (Ellies et al., 1997). The homeodomain is a compact structure consisting of three α-helices; however, the rest is not predicted to form a secondary structure. The typical 3D structure of proteins is due to the specific amino acid sequence and various forces, including Van der Waals and electrostatic forces (Habchi et al., 2014; Tompa & Fuxreiter, 2008). The loss of these protein-binding forces compels a protein into an intrinsically disordered protein (IDP) state (Tompa & Fuxreiter, 2008). The resultant protein resembles a coil shape instead of the typical α-helix and β-sheet structures (DeForte & Uversky, 2017). Proteins can contain both folded structures and disordered regions. The primary amino acid sequence determines the folding or non-folding ability of a protein (Dunker et al., 1998). For instance, Trp, Cys, Tyr, Ile, Phe, Val, Asn, and Leu are order-promoting residues, whereas Arg, Pro, Gln, Gly, Glu, Ser, Ala, and Lys are considered disordered-promoting residues in protein structure (O'Shea et al., 2017).

1.5.7 Homeodomain (HD) proteins and their target DNA recognition

Many HD proteins have little DNA-binding specificity as they recognize a core consensus binding motif of TAAT. The flanking sequence provides additional specificity and the binding preferences of some homeoprotein families can be distinguished in this way. In general, DNA recognition by TFs is based on two mechanisms: base readout and shape readout (Dror et al., 2014). The minor groove contributes to binding strength by the N-terminal arm of the homeodomain. The minor groove width overall increases the negative electrostatic potential. As a result, positively charged amino acids such as arginine, lysine, and histidine are attracted, which change the overall geometry of the DNA. DNA 3D structure and the local DNA sequence have a high correlation. Different amino acids can produce the same DNA structure, but a single variation in the sequence can also change
the overall DNA landscape. In the N-terminal arm of the homeodomain, amino acid positions 2-7 are crucial for shape readout (Dror et al., 2014). The HD recognizes the AT-rich sequence by interacting with the DNA major groove in accordance with the HD-TF base readout mechanism (Rohs et al., 2009). Amino acid 50 of the homeodomain is believed to play a vital role in target recognition, and Dlx5 Q50 and N51 play an indispensable role in DNA binding (Duan & Nilsson, 2002). Furthermore, a mutation in the Q186 amino acid in the HD to histidine region reduces Dlx5 binding affinity and specificity (Proudfoot et al., 2016).

HD proteins also increase both DNA binding affinity and specificity by binding with cofactors, interactions with flanking regions of DNA or through post-translational modifications which augment their protein-protein interactions (Yu et al., 2001). For example, the zinc finger-like homeodomain, POU-homeodomain, and paired-like homeodomain proteins have additional DNA binding domains that recognize additional flanking motifs (Bürglin & Affolter, 2016). Multiple HDs are present in a single protein for enhanced DNA binding specificity in a few cases. An extreme case can be seen in the C. elegans homeobox protein, ceh-100, where twelve tightly packed HDs are present (Hench et al., 2015). In Dlx proteins, two tryptophan residues found close to the C-terminal end of the homeodomain may facilitate their interaction with PBS-family cofactors, such as Pbx and extradenticle (Exd) (Chang et al., 1995; Neuteboom et al., 1995). Experimentally, it was demonstrated that Dlx2 and Dlx5 can form a heterodimer complex with mesodermal proteins Msx1 and Msx2 (Zhang et al., 1997). In this complex, the homeodomain of both proteins interacts, preventing DNA binding of either protein.

### 1.5.8 Dlx gene targets

Being a family of TFs, Dlx genes control the activation and repression of many genes (Lindtner et al., 2019). Regarding what is known about Dlx targets, it is interesting to note that some Dlx genes can activate other members of the same gene family. For example, Dlx1 and Dlx2 can activate transcription through a Dlx5/Dlx6 enhancer in cultured cells (Yu et al., 2001). It is shown that there is a single binding site for Dlx2 on the Wnt-1 locus referred to as HBS-1 (Zeng et al., 2020). In addition, BMP-2 activates Dlx2, which induces
the differentiation of chondroblast (Piper et al., 1999; Xu et al., 2001). Some Dlx genes, such as Dlx1, Dlx2 and Dlx5, have the ability to regulate the glutamic acid decarboxylase gene (Zerucha et al., 2000). This event is vital in the differentiation of GABAergic neurons. It is known that Arx is primarily involved in the migration of GABAergic cells, and mutation of Arx orthologue can cause several neurodegenerative disorders. Colasante et al. (2008) demonstrated that Dlx2 is directly involved in regulating the Arx gene, which indirectly affects the differentiation of GABAergic neurons. To understand the molecular relationship of these genes, bioinformatics searches and experimental results showed that TAAT, which was found within the regulatory elements of Arx, is the sequence required for Dlx2 to bind Arx locus and mediate Arx activity (Colasante et al., 2008). Dlx3 has been implicated in activating human chorionic gonadotropin (a heterodimeric glycoprotein) in the placenta. One of the targets of Dlx3 is the gonadotropin α subunit, common to all members of the glycoprotein family (Roberson et al., 2001). Dlx4 has the ability to act as both an activator and repressor (Fu et al., 2001). Dlx4 can activate MYC and globulin transcription factor 1 (GATA1) (Merlo et al., 2000; Shimamoto et al., 1997); moreover, GATA1 is involved in red blood cell maturation. The ectopic expression of Dlx4 can repress apoptosis by regulating intercellular adhesion molecule 1 in hematopoietic cells.

There are three isoforms of Dlx4, and all isoforms encode a homeodomain protein that can bind to the identical sequence for gene activation (Panganiban & Rubenstein, 2002). The upregulated expression of Dlx4 is found in some leukemias, whereas down-regulation is related to hematopoietic stem cell differentiation (Chase et al., 2002). One of the isoforms of Dlx4 is beta protein 1 (BP1) which was found to repress the expression of the β globin protein (Chase et al., 2002).

Dlx5 regulates many developmentally expressed genes, such as Runx2, Gbx2, Notch1/3 and Irs2. Lee et al. (2005) showed that Dlx5 mediates the BMP2-induced differentiation of C2C12 cells by, in part, activating Runx2, (Lee et al., 2005), osteocalcin (Newberry et al., 1998; Ryoo et al., 1997) and alkaline phosphatase (Kim et al., 2004). Promoters of
these genes harbour a short homeodomain response element (HDRE) in the form of TAAT or ATTA (Sajan et al., 2011). In a site-directed mutagenesis experiment, it was discovered that, again, a TAAT site is necessary for Dlx5 binding and subsequent regulation of Runx2 (Lee et al., 2005). Immunoprecipitation of chromatin (ChIP) illustrated the direct binding of Dlx5 as a TF with this motif in the promoters of Bmper, Embf1 and Msx1 genes (Sajan et al., 2011). Sp7, a key TF in osteoblasts, requires Dlx5 for recruitment to the regulatory regions of osteoblast-specific genes (Hojo et al., 2016). Chip-seq. data for chromatin binding by Sp7 identified the TAATT site at the center of binding peaks, rather than the GC-rich binding motif expected for this zinc-finger TF. Sp7 and Dlx5 interact through the zinc-finger domain. So, Dlx5 helps anchor Sp7 protein to its target DNA (Hojo et al., 2016).

Down-regulation of Gbx2 and hepatocyte growth factor (Hgf) are related to the development of shorter otic capsules, reduced mandible, and impaired development of motor neurons (Byrd & Meyers, 2005). Gbx2 plays a pivotal role in forming rhombomeres in the hindbrain, and this essential developmental protein production is activated by Dlx5 (Jeong et al., 2008). Lastly, it has been determined that in T-cell lymphoma, Notch1/3 and Irs2 are additional direct transcriptional targets of Dlx5 (Tan et al., 2017).

Dlx proteins undergo post-translational modifications, which increase or decrease their transcriptional potential. In osteoblast differentiation, Dlx5 is phosphorylated by calmodulin-dependent kinase II, Protein kinase A and Protein kinase B, enhancing its transactivational potential (Jeong et al., 2011; Seo et al., 2009). In C2C12 cells, p38 phosphorylates Dlx5 at residues, Ser34 and Ser217, leading to enhanced Osterix (Osx) gene expression (Ulsamer et al., 2008). Contrastingly, in developing keratinocytes, the Dlx3 homeodomain is phosphorylated by protein kinase C, and this calcium-dependent phosphorylation reduces the binding affinity of Dlx3 with DNA (Park et al., 2001). Certain amino acids in Dlx proteins facilitate essential functions and hence are conserved. As an example, transcriptional activation is stimulated by proline-enriched domains located at amino and carboxy terminals (Panganiban & Rubenstein, 2002). Homeodomain proteins’
N-terminal and C-terminal domains are proline-rich and facilitate protein oligomerization and transcriptional activation (Mermod et al., 1989; Tanaka & Herr, 1990; Xiao et al., 2000). To date, however, it has not been shown how Dlx proteins mechanistically stimulate RNA polymerase activity.

1.6 Rational and thesis objectives

Examination of the expression of multiple genes in the PA1 of $Dlx5/6^{-/-}$ embryos has provided further insight into $Dlx5/6$-mediated gene expression. Whether the involvement of Dlx proteins in regulating these genes is direct or indirect has not yet been demonstrated. Collectively, Dlx5 and Dlx6 are clearly involved in regulating the fate of the lower jaw and are important effectors of the intricate signalling pathways involved in jaw morphogenesis. Most of our understanding of jaw patterning comes from genetic manipulations in mice and zebrafish. Little is known about the gene regulatory network in avian embryos and to what extent regulatory connections are conserved. This thesis focuses on the Dlx-mediated regulation of genes in the first PA of chick embryos. In this thesis, I documented the expression of all Dlx genes in the pharyngeal arches of chicken embryos and examined their dependence on the endothelin signalling pathway. Some candidate Dlx target genes – Pitx1, Gsc, Gbx2, Hand1 and Hand2 – whose expression is lost or down-regulated in Dlx5/6 null mouse embryos were screened for Dlx5-responsiveness in chicken embryos. The genomic loci of these candidate Dlx-regulated genes were examined using bioinformatics tools to identify nearby sequence-conserved regulatory regions. To identify novel Dlx targets, we examined the intersection of two publicly available datasets: p300 ChIP-seq data from E13.5 mouse craniofacial tissue (FaceBase) and first arch-active enhancer regions from transgenic mouse embryo screens (VISTA). The orthologous chicken loci for both candidate and novel downstream genes were screened for activity in migratory CNCCs or pharyngeal arch ectomesenchyme.
1.7 Hypothesis

It is hypothesized that Dlx5 and Dlx6 contribute to jaw patterning by directly regulating genes in the mandibular branch of the first pharyngeal arch via TAATT-containing *cis*-regulatory regions that are conserved between mammals and birds (Fig 1.3).

![Diagram](image)

**Fig. 1.3:** Possible modes of Dlx-mediated gene regulation downstream of Edn signalling in chicken PA1

Dlx5/6 is regulated by endothelin signalling; Dlx5/6 regulates the expression of some genes directly or through other Dlx paralogues or maybe by regulation of some other unknown genes.
1.8 Thesis Organization and Author Contribution

The endeavour and results of this work are organized as follows:

**Chapter 2**: All materials and methods used in the work are described

**Chapter 3**: *Dlx* gene expression in the developing chick pharyngeal arches and dependence on endothelin signalling. **Chapter 3** is submitted for publication as: **Sohail, A** and Bendall, AJ.

**Chapter 4**: Putative targets of Dlx5/6 and their direct or indirect correlation with the endothelin signalling pathway. **Chapter 4** represents a manuscript in preparation for later submission as: **Sohail, A** and Bendall, AJ.

**Chapter 5**: Enhancer Prediction and validation in chick embryos by microinjections and electroporation. The Dlx5 and Dlx6 responsiveness of these predicted putative enhancer elements are also validated by reporter assay. **Chapter 5** represents a manuscript in preparation for later submission as: **Sohail, A**; *Walchyshyn, R*; *Nicoll, O*; and Bendall, AJ.

**Chapter 6**: General discussion, significance, and work in progress which are regarded as future directions.
Chapter 2
Material & Methods

2.1 Embryos and staging

All experiments using chicken embryos were performed following the University of Guelph's institutional regulations, in accordance with the Canadian Council on Animal Care (CCAC). Fertilized chicken eggs (Plymouth Barred Rock) were obtained from Arkell Poultry Research Facility at the University of Guelph, and specific pathogen-free (SPF) eggs (White Leghorn) were purchased from Charles River, Ottawa (Canada). All eggs were incubated at 38°C in a humidified chamber. The embryos were stage-matched using several morphological criteria such as overall size, limb bud shape, and somite number (Hamburger & Hamilton, 1951). Older embryos were also aged in embryonic days (E), corresponding to incubation time.

2.1.1 Cell lines and transfection

The chicken fibroblast cell line, DF-1 (ATCC CRL-12203) and human embryonic kidney 293T cells (HEK 293T) were cultured and maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (FBS), 100 U/ml penicillin, 100 µg/mL streptomycin and 2 mM L-glutamine. Cell cultures were maintained in 5% CO₂ at 37°C. Cells were seeded at 60-70% confluency and transfected the following day with polyethyleneimine (PEI) at 4.5 µg PEI/µg of plasmid DNA. The media was changed 4 hr post-transfection.

2.2 Extraction of high molecular weight genomic DNA

Fertilized chicken eggs were incubated until embryonic day 10 (E10), and 1-2 g of the liver was snap-frozen in liquid nitrogen and manually ground to a fine powder with a chilled
mortar and pestle. Liquid nitrogen was allowed to evaporate then the powdered tissue was mixed with 4 ml ice-cold phosphate buffer saline (PBS). The tissue was digested overnight (18 hr) at 50°C in 20 ml lysis solution (10 mM Tris-HCl pH 7.5, 2 mM ethylenediaminetetraacetic acid (EDTA), 1% sodium dodecyl sulphate (SDS), 0.5 mg/ml Proteinase K). The next day, the sample was extracted with an equal volume of 50:49:1 phenol:chloroform:isoamyl alcohol with rocking for 30 minutes at room temperature (RT). Following centrifugation at 4,000 rpm in a benchtop centrifuge, the aqueous phase was retrieved and centrifuged in a high-speed centrifuge at 15,000 rpm for 15 minutes. Genomic DNA was precipitated with 0.8 vol. isopropanol and the pellet was washed with 70% ethanol and then 100% ethanol. The air-dried pellet was resuspended at 1 mg/ml in TE (10mM Tris-HCl, 1mM EDTA, pH 8.0) with heating to 65 °C for several hours.

2.3 Embryo powder

Chicken embryos at E10 were collected and homogenized with a mechanical homogenizer in a minimal volume of ice-cold PBS. Four volumes of ice-cold acetone were added, and the mixture was incubated on ice for 30 minutes. The mixture was transferred to Corex tubes and centrifuged at 10,000 xg for 10 minutes at 4°C. The supernatant was discarded, the pellet washed with ice-cold acetone, and re-centrifuged. The pellet was air-dried overnight (16-18 hours) and ground to a fine powder with mortar and pestle. Embryo powder was stored long-term at -20°C.

2.4 RNA extraction

The mandibular and hyoid arches (total of 30 mg wet tissue) were dissected from 2-3 embryos at HH25. Total RNA was extracted with an RNeasy Minikit according to the manufacturer’s instructions. Genomic DNA contamination was removed using a RapidOut DNA Removal Kit. RNA integrity was confirmed by separating 1-2 μl of RNA on a 1% TAE agarose gel on ice for 15-20 minutes at 150 V and imaging with Red Safe dye using a Gel-Doc (Bio-Rad). The RNA concentration was determined using a NanoDrop 8000 (ThermoScientific).
2.5 Reverse transcription and preparative polymerase chain reaction (PCR)

First-strand cDNA was synthesized using qScript cDNA supermix (Quantabio). One microgram of total RNA was converted to first-strand cDNA using 4 µl of qScript mix in a total volume of 20 µl. The reaction was carried out as follows: 5 minutes at 25°C, 30 minutes at 42°C and 5 minutes at 85°C. PCR reactions contained: 1µl cDNA, 20 pmol each forward and reverse primer, 200 mM dNTPs, 2 U HF-Phusion DNA polymerase, and 1x Phusion buffer in 50 µl. PCR was performed using the conditions shown in Table 2.1. All primers were designed against gene sequences deposited in The National Center for Biotechnology Information (NCBI).

Table 2.1: PCR cycling conditions

<table>
<thead>
<tr>
<th>Cycle No.</th>
<th>Denature</th>
<th>Anneal</th>
<th>Extend</th>
<th>Hold</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>95°C, 5 min</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-31</td>
<td>95°C, 30 sec*</td>
<td>55°C, 30 sec*</td>
<td>72°C, 2 min</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td></td>
<td></td>
<td>72°C, 10 min</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td></td>
<td></td>
<td></td>
<td>4°C hold</td>
</tr>
</tbody>
</table>

PCR products were separated on 1.2% TAE agarose gels containing Red Safe to confirm amplicon size and the PCR products were gel-extracted using a Qiaex II kit (Qiagen). Cloning details are listed in Table 2.2. Restriction-digested PCR products were treated with Strata Clean Resin (Stratagene) to remove restriction enzymes prior to ligation.

2.6 Plasmid cloning

Plasmids and ligation mixtures were transformed into chemically competent DH5α E. coli by heat shock at 42°C for 45 seconds. Plasmids were maintained under selection (100
µg/ml Ampicillin or 100 µg/ml Kanamycin) in liquid Luria Bertani (LB) or solid (LB-agar) cultures. Plasmids were recovered from overnight cultures (3 ml miniprep or 50-100 ml midiprep) by alkali lysis using a variety of commercial midiprep kits.

T4 DNA ligase (5 Weiss units/ µl) was used to ligate sequences with sticky ends, either overnight at 16°C or for 2 hours at RT. Blunt-ended PCR products were cloned with a TOPO blunt end kit (Invitrogen). TOPO cloning was achieved by incubating 1 µl of TOPO vector, 2 µl of TOPO salt solution and 2 µl of PCR product on the ice for 30 minutes. Transformation of TOPO reaction products used Super Optimal broth with Catabolite repression (SOC) in place of LB for higher plasmid transformation efficiencies. Recombinant plasmids were screened by restriction digestion of miniprep DNA or colony PCR using insert-specific primers.

For miniprep digests, antibiotic-resistant colonies were transferred to 3 ml of LB + 100 µg/ml Ampicillin and cultured overnight at 37°C with shaking at 220 rpm. Cultures were centrifuged for 3 minutes at 14,000 xg. Pellets were resuspended in 100 µl PEB1 (50mM glucose, 25mM Tris-HCl, 10 mM EDTA, pH8) and incubated at RT for 5 minutes. Subsequently, 200 µl of PEB2 (0.2 NaOH, 1% (w/v) SDS) solution was added, and samples were incubated on ice for a further 5 minutes before adding 150 µl PEB3 (3 M KOAc, 2 M acetic acid) with another 5 minutes on ice. Samples were centrifuged for 1 minute, and 1 ml of ethanol was added to the supernatant, followed by another 5 minutes of centrifugation at 14,000 xg and 4°C. Plasmid DNA was resuspended in 50 µl TE buffer containing 50 µg/ml of RNase and incubated at 37°C for 5 minutes. DNA was extracted with an equal volume of 50:49:1 phenol:chloroform: isoamyl alcohol and
<table>
<thead>
<tr>
<th>Gene Name Accession No.</th>
<th>Primers*</th>
<th>Amp.size (bp)</th>
<th>Cloning Site</th>
<th>Vector</th>
<th>RE† &amp; RNA pol.</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dlx1 NM_001045842</td>
<td>For: TCT CCT TCT CCC ATG TCC CA Rev: TTC GGG CTC ACA TAA GCT GC</td>
<td>685</td>
<td>ECORI</td>
<td>TOPO</td>
<td>Xhol- Sp6</td>
<td>This study</td>
</tr>
<tr>
<td>Dlx2 XM_025152605</td>
<td>-----------</td>
<td>610</td>
<td>------</td>
<td>Blue Script</td>
<td>EcoRI-T7</td>
<td>Puelles et al. (2000)</td>
</tr>
<tr>
<td>Dlx3 NM_204804</td>
<td>For: GAT CTC ATG AGC GGC TCC TCC TGC AAG Rev: TAT AGG ATC CCC GTA AAC GGC GCC CGG GTT</td>
<td>833</td>
<td>Xbal-BamHI</td>
<td>Blue Scr</td>
<td>Xbal –T7</td>
<td>Zhu &amp; Bendall(2006)</td>
</tr>
<tr>
<td>Dlx4 XM_015299468</td>
<td>For: GTC AAG ATC TGG TCC CAG AAC Rev: GCA GGT ACC GAG GAG CAG GAG CAG</td>
<td>635</td>
<td>ECORI</td>
<td>TOPO</td>
<td>Xhol-Sp6</td>
<td>This study</td>
</tr>
<tr>
<td>Dlx5 NM_204159</td>
<td>For: T A TA GGA TCC ATG ACA GCA GTG TTT GAC Rev: T A TA TCT AGA TCA GAG GTG CTT CCC GGA G</td>
<td>861</td>
<td>BamHI-HindIII</td>
<td>Blue Script</td>
<td>BamHI- T7</td>
<td>Bendall et al. (2003)</td>
</tr>
<tr>
<td>Dlx6 NM_001080890</td>
<td>For: G ATC AAG CCT GTG TGG CAG TCG GCA ACC Rev: T A TA GGA TCC CAA CCC CCA CGA GAG CGA</td>
<td>804</td>
<td>BamHI-HindIII</td>
<td>pGEM11Zf</td>
<td>BamHI- Sp6</td>
<td>Quach et al. (2016)</td>
</tr>
<tr>
<td>Gsc NM_205331</td>
<td>For: T A TA GGA TCC CCT GCG AGCATG TGG AGC Rev: T A TA TCT AGA TCA GCT GGA GTC CAA</td>
<td>656</td>
<td>ECORI</td>
<td>TOPO</td>
<td>Xhol-Sp6</td>
<td>This study</td>
</tr>
<tr>
<td>Pitx1 NM_00116768</td>
<td>For: TATA GATCC GAT TCC TTT AAA GGT GGA ATG Rev: TATA TCTAGA TGT GTC GGA GGT GTC TCG</td>
<td>935</td>
<td>BamHI-Xbal</td>
<td>Blue Script</td>
<td>BamHI- T7</td>
<td>This study</td>
</tr>
<tr>
<td>Hand1 NM_204965</td>
<td>For: TATA GATCC GGG TAC CAG CAC CAC CAC Rev: T A TA TCTAGA C TCA GGG GTT CAG TCC CAG A</td>
<td>606</td>
<td>ECORI</td>
<td>TOPO</td>
<td>BamHI- T7</td>
<td>This study</td>
</tr>
<tr>
<td>Hand2 NM_204966.3</td>
<td>For: T A TA GGA TCC AGT CTT GTG GGC GGC GCC TCC Rev: T A TA TCT AGA TCA TCG CTT GAG CTC CAG CG</td>
<td>650</td>
<td>ECORI</td>
<td>TOPO</td>
<td>Xhol- Sp6</td>
<td>This study</td>
</tr>
<tr>
<td>Gbx2 NM_205068.2</td>
<td>For: ATGGACAGCGACCTCGA CTA Rev: TCGCTCGTTTGCTTTTCAG</td>
<td>523</td>
<td>ECORI</td>
<td>TOPO</td>
<td>Xhol- Sp6</td>
<td>This study</td>
</tr>
</tbody>
</table>

All primers were synthesized by Integrated DNA Technology (IDT) and are shown as 5' → 3'.

†Restriction enzyme used to linearize plasmid template for antisense transcription.
precipitated with 0.1 volume 0.3 M NaOAc pH 5.2 and 2.5 volume 70% ethanol. Ten percent of the recovered plasmid was digested with the appropriate restriction enzyme combination to screen for the desired insert.

A colony was streaked onto a selective LB-agar plate and then dipped into a 25 µl PCR mixture with vigorous mixing for colony PCR screening. PCR for colony screening followed the conditions above, including the cycling conditions in Table 2.1. All cloned products were verified by Sanger sequencing at the Advanced Analysis Center Genomics Facility, University of Guelph.

### 2.7 Riboprobe synthesis

Ten micrograms of plasmid were linearized with the restriction enzyme shown in Table 2.2, extracted with 2 volumes of Phenol/chloroform and ethanol precipitated. The pellet was resuspended at 0.5 µg/µl in 20 µl of RNase-free water. Antisense riboprobes were transcribed from a 1 µg linearized DNA template in a 20 µl reaction using a digoxigenin (DIG) labelling mix (Roche), 20 U rRNasin, and 20 U RNA polymerase (T7 or SP6) for 3 hours at 37°C. Ten percent of the reaction was separated on a 1% TAE agarose gel to confirm transcriptional efficiency. To the remaining mixture, 20 U rRNasin and 4 U TURBO RNase-free DNase were added and incubated for an additional 15 minutes at 37°C. RNA was precipitated overnight at -20°C with 0.4 M lithium chloride (LiCl), 40 mM EDTA, and 75% ethanol. Immediately before use, RNA was centrifuged for 15 minutes at 14,000 xg at 4°C, washed with 70% ethanol, air-dried, resuspended in 50 µl of RNase-free water and immediately added to the hybridization buffer (Zhu & Bendall, 2006).

### 2.8 Whole-mount In situ hybridization (WMISH)

Embryos between HH16 and HH25 were fixed overnight (16-18 hours) at 4°C in 4% (w/v) paraformaldehyde (PFA) in PBS (1.37 M NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄·7H₂O, 1.8 mM KH₂PO₄). Embryos were dehydrated in a 25%, 50%, and 75% methanol series in PBT (PBS containing 0.1% Tween-20), finished in 100% methanol and stored in fresh
100% methanol. On day one, embryos were rehydrated in a reverse methanol series (75%, 50%, 25%) to PBT, bleached with 6% hydrogen peroxide for 1 hour at room temperature with continuous rocking, then permeabilized with Proteinase K in PBS at room temperature as outlined in Table 2.3

Table 2.3: Proteinase K digestion conditions for chick embryos

<table>
<thead>
<tr>
<th>HH Stage</th>
<th>Time (min)</th>
<th>Prot. K (µg/ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-16</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>17-23</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>24-26</td>
<td>10</td>
<td>5</td>
</tr>
</tbody>
</table>

Digestion was stopped by washing the embryos for 10 minutes with 2 mg/ml glycine in PBT, and the embryos were post-fixed in 4% PFA and 0.2% glutaraldehyde in PBT for 20 minutes. Embryos were washed 3x 5 minutes with PBT. Fixed embryos were incubated with pre-hybridization solution (50% formamide, 0.75 M NaCl, 75 mM trisodium citrate.2H₂O, 1% SDS, 50 µg/ml yeast tRNA, 50 µg/ml Heparin, pH 4.5) for 1 hour in a 70°C water bath. DIG-labelled antisense probes were added to a final concentration of 1 µg/ml, and embryos were incubated overnight at 70°C. At least 4 embryos were incubated with each probe for each stage. Probes were recovered and re-used up to five times. Post-hybridization, embryos were washed 3x 30 minutes with the solution I (50% formamide, 0.6 M NaCl, 60 mM trisodium citrate.2H₂O, pH 4.5) at 72°C. Subsequently, embryos were washed 3x 5 minutes with solution II (50% formamide, 0.3 M NaCl, 30 mM trisodium citrate.2H₂O, pH 4.5) at 65°C. Embryos were washed 3x 5 minutes in TBST (137 mM NaCl, 2.7 mM KCl, 25 mM Tris-HCl, pH 7.5, 0.1% Tween-20, 2 mM levamisole) then blocked with 10% horse serum in TBST at room temperature for 2.5 hours with continuous shaking.

Alkaline phosphatase-conjugated anti-digoxigenin antibody Fab fragments (Roche Cat # 11093274910) were pre-absorbed with 3 mg/ml chick embryo powder and 1% horse
serum in TBST for one hour at 4°C. Antibody was centrifuged at 4,000 xg at 4°C, and the supernatant was diluted four-fold with 1% horse serum in TBST. Embryos were immersed in the antibody solution overnight (16-18 hours) at 4°C with gentle rocking. Embryos were washed with TBST 6x 90 minutes at RT, then overnight at 4°C. On day four, embryos were washed 3x 10 minutes with NTMT (100 mM Tris-HCl, 100 mM NaCl, 50 mM MgCl$_2$, 0.1% Tween-20, 2 mM levamisole, pH9.5). Riboprobes were detected in the dark at RT with 250 µg/ml NBT and 130 µg/ml BCIP in NTMT (Zhu & Bendall, 2006). The reaction was stopped with 1 mM EDTA in PBS, pH 5.5. Embryos were dehydrated through a methanol series, 15 minutes for each wash, and washed extensively with PBT (pH 7.5) for several days to minimize background staining. Embryos were photographed using a light microscope (Zeiss) equipped with the Q capture software.

2.9 Bosentan treatment

Bosentan was added to corn oil at 5 mg/ml, incubated for 30 minutes at 37°C, then sonicated till the solution was clear. Eggs were incubated narrow end down for 60 hours to reach HH13. A window was made in the shell, and 60 µl of the drug or solvent was pipetted directly onto the shell membrane. Eggs were sealed and re-incubated for an additional 24-48 hours (corresponding to HH17-23) for in situ hybridization (ISH) or qPCR or up to E11 for skull staining (Kempf et al., 1998).

2.10 Quantitative polymerase chain reaction (qPCR)

RNA was extracted from control or Bosentan-treated HH22-23 embryos using an RNeasy Minikit per the manufacturer’s instructions. Total RNA was extracted from 30 mg of homogenized tissue (mandibular and hyoid arches dissected and pooled from 2-3 embryos), and genomic DNA contamination was removed using a RapidOut DNA Removal Kit. One microgram of RNA was converted to first-strand cDNA using the qScript superMix. RNA integrity was confirmed by running 1-2 µl of RNA on 1% TAE agarose gel containing Red Safe and was imaged using a Gel-Doc (BioRad). The RNA concentration was determined using a NanoDrop 8000 (ThermoScientific). All gene expression analysis was done with SYBR Green on a StepOnePlus real-time PCR thermocycler (Applied
Biosystems) in the Advanced Analysis Center Genomics Facility, University of Guelph. Reagents and their final concentrations are: 1x SYBR Green PCR master mix (7.5 μl), 0.2 μM of each forward and reverse primer, and 5 μl of 5X diluted cDNA. Sterile water was used to make the reaction volume 15 μl. The amplification program consisted of 10 minutes at 95°C, followed by 40 cycles of 95°C for 15 seconds, 60°C for 15 seconds and 72°C for 45 seconds. A no-template control was also run for each set of genes. The expression of the housekeeping gene, Succinate Dehydrogenase (SDHA), was used to normalize the data. Primer efficiencies were determined using standard curves with serial dilutions of cDNA. Values presented are means of at least three biological replicates, each performed with three technical replicates.

The gene expression level was analyzed using the $2^{-\Delta\Delta C_T}$ method of Livak & Schmittgen (Livak & Schmittgen, 2001). First, we normalized the $C_T$ of genes of interest (GOI) to the reference gene with both Bosentan treated and control embryos.

$\Delta C_T$ (GOI, Bosentan-treated PA) = $C_T$ (GOI, Bosentan-treated PA) - $C_T$ (ref. gene, Bosentan-treated PA) and

$\Delta C_T$ (GOI, control PA) = $C_T$ (GOI, control PA) - $C_T$ (ref. gene, control PA).

Normalization was done using the following equation:

$\Delta\Delta C_T = \Delta C_T$ (target genes) - $\Delta C_T$ (reference gene; SDH)

Finally, expression ratios were calculated using:

$2^{\Delta\Delta C_T}$ = Normalized expression ratio

All primer sequences with their resultant amplicon sizes are tabulated below.
Table 2.4: qPCR primers and their respective amplicon size

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Primers (5’→ 3’)</th>
<th>Amplicon size (bp)</th>
<th>Primer efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dlx1</td>
<td>For: ATGCCTGAGAGTCTAAACAGC Rev: AACAGTGCGTGAATAGTGTC</td>
<td>118</td>
<td>100.8</td>
</tr>
<tr>
<td>NM_001045842.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dlx3</td>
<td>For: AGCCCAATACCTGGCGCTG Rev: CGGCACCTCGCCGTTCTTG</td>
<td>133</td>
<td>99.8</td>
</tr>
<tr>
<td>NM_204804.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dlx4</td>
<td>For: TCCCCATGTGTGACAGCACCA Rev: CATCTCCATTGCTGGGTTG</td>
<td>123</td>
<td>98.6</td>
</tr>
<tr>
<td>XM_015299468.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dlx5</td>
<td>For: CACCGAGGAGGAGCCTA Rev: ATAAATAGTGCCGTTTGC</td>
<td>129</td>
<td>100.2</td>
</tr>
<tr>
<td>NM_204159</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dlx6</td>
<td>For: CGGCACCCGGGCTCTCA Rev: CGTCTTTCGATCCGGTCGTTT</td>
<td>161</td>
<td>103.4</td>
</tr>
<tr>
<td>NM_001080890</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hand1</td>
<td>For: AGGAGGAGAACGGAGAGCATC Rev: CAGGTAGGCCGATGCTGGG</td>
<td>129</td>
<td>97.6</td>
</tr>
<tr>
<td>NM_204965.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hand2</td>
<td>For: GCAGGACTCAGAGCATCAACA Rev: AGGTAGGCGATGTCTGGTG</td>
<td>124</td>
<td>102.4</td>
</tr>
<tr>
<td>NM_204966.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gbx2</td>
<td>For: AGCTGAGCGAGGTGAGGTGA Rev: GGATGGGCACGACGATCTGG</td>
<td>129</td>
<td>98.1</td>
</tr>
<tr>
<td>NM_205068.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pitx1</td>
<td>For: TCTTTAAAGGTGGAATGAAT Rev: GATTCACCTGCGAGTTTCTC</td>
<td>137</td>
<td>100.8</td>
</tr>
<tr>
<td>NM_001167686.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gsc</td>
<td>For: TAGTCCGGTGCGCTTTTGG Rev: CGTCTCGTCTGTCTGTTCTGT</td>
<td>122</td>
<td>98.9</td>
</tr>
<tr>
<td>NM_205331</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDHA</td>
<td>For: CAGGGATGTAGTGTCTGTG Rev: GGAATAGGGCTCTTTAGTG</td>
<td>186</td>
<td>98.9</td>
</tr>
<tr>
<td>NM_001277398.1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.11 Skull staining

E11 embryo heads were fixed in 95% ethanol for 3-4 days and then immersed overnight in 0.03% Alcian blue and 3.48 M acetic acid in 80% ethanol to stain cartilage. Embryos were washed with 95% ethanol for 3 hours, then cleared with 2% potassium hydroxide (KOH). Embryos were then immersed in 0.3% Alizarin Red in 1% KOH overnight (16 hours) to stain bone. Skulls were washed with 1% KOH/20% glycerol and stored in 1:1 glycerol:ethanol.
2.12 Enhancer predictions

All genomic regions used in this study were downloaded from VISTA Enhancer Database (https://enhancer.lbl.gov) unless otherwise stated. The Hand2 enhancer was identified in Charité et al. (2001). The Gsc sequence was predicted in this study based on the histone acetylation mark (H3K27ac) and the level of sequence conservation among a variety of species.

Previously unrecognized chicken pharyngeal arch enhancers were first identified in the overlap of two publicly available murine data sets: FaceBase p300 ChIP-seq peak sequences and VISTA Enhancer-validated enhancer sequences mapped to the Mouse GRCm38/mm10 genome. A VISTA locus is considered to be an enhancer if it showed reproducible lacZ reporter expression within a common tissue in at least 3 different embryos. Experimental data for all verified enhancer elements within the VISTA Enhancer Browser is accessible through https://enhancer.lbl.gov/. Raw data and detailed FaceBase protocol are available at https://www.facebase.org/. Both datasets were also available for view and use in the UCSC Genome Browser (https://genome.ucsc.edu/). The orthologous chicken genomic sequence was identified from the Chicken GRCg6a/galGal6 genome using UCSC Genome Browser's BLAT tool, and sequence conservation between the two sequences was analyzed through “glocal” alignment (rVISTA) and multi-species alignment (phastCon, phyloP – UCSC Genome Browser) algorithms. rVISTA “glocal” alignment conditions and calculations are available here: http://genome.lbl.gov/vista/index.shtml. Loci with at least 80% sequence identity between mouse and chicken over a minimum of 100bp were further validated.

The gene expression profile for all genes within 500 kb of each conserved sequence was examined between E9.0 and E12.5 using the Jackson Laboratory’s MGI Mouse Expression Database (http://www.informatics.jax.org/expression.shtml) to identify at least one gene for each enhancer with verified expression in the PA1. GEISHA was used to examine the expression of the chick orthologue (http://geisha.arizona.edu/geisha/). All sequences that fulfilled all of the above criteria were considered candidate enhancer
regions and a subset was cloned into ptk-EGFPv2 (Uchikawa, 2008) for in ovo activity assays.

2.13 Microinjection & electroporation

Fertilized Plymouth Barred Rock eggs were obtained from the Arkell Poultry Research, University of Guelph and incubated at 38°C for 44-48 hours to reach HH9-10. Approximately 3 ml of albumin was removed, the egg was windowed, and the vitelline membrane was removed from the anterior end of the embryo. Reporter plasmid ptk-EGFPv2, encoding enhanced GFP under the control of the HSV1 thymidine kinase promoter (Uchikawa, 2008), was mixed 1:1 with a control pCAGGS-mCherry plasmid (Hitoshi Niwa et al., 1991) (encoding a membrane-localized cherry fluorescent protein with constitutive expression from the chick β-actin promoter and CMV enhancer) to a final concentration of 2 µg/µl each, and 0.1 volume of 1% Fast Green dye was added to visualize the injected volume. Plasmids were loaded into a heat-pulled Borosilicate glass capillary tube (1 mm outer diameter x 0.75 mm inner diameter). Approximately 0.2 µl was injected into the hindbrain lumen, aided with Leica MZ 75 microscope, to target the migratory NCCs. Gold-plated electrodes (0.8 mm diameter) were placed on either side of the embryos, 4 mm apart, at the hind- and mid-brain level. Embryos were covered with 100 µl of sterile PBS immediately prior to electroporation with a BTX Electro Square Porator ECM 830. Electroporation conditions were three 50 ms pulses at 15 V with an inter-pulse interval of 0.1 sec. Embryos and electrodes were manipulated under magnification using a Leica MZ75 microscope. Shell windows were sealed with clear tape, and electroporated embryos were re-incubated for 24 or 48 hours. Live embryos were collected and washed in PBS, and all membranes were dissected before viewing on a Nikon Eclipse Ti2 microscope. Images were collected with NIS Elements software.

2.14 Transcription assay

For luciferase assays, HEK293T cells were seeded at 1.2x10⁵ cells per well in 12-well plates. The cells were transfected the following day with variable amounts (20 to 320 ng) of pcDNA3 plasmids containing Dlx5 and Dlx6 coding sequences, 16 ng of pRL-SV40-
Renilla, and 400 ng of ptk-luc, containing the predicted enhancer regions (section 2.12), after 24 hr of seeding. The total pcDNA3 backbone levels were kept consistent by including empty pcDNA3. Twenty-four hours post-transfection, the cells were rinsed in PBS and then rocked manually for 15 minutes in 200 μl of Passive Lysis Buffer (Biotium). Lysed cells were scraped to collect, and the contents were transferred into Eppendorf tubes and centrifuged at 14,000 rpm for 3 minutes. 10 μl of each sample was loaded into each well in a 96-well plate, with three technical replicates per sample. Six microlitres of D-Luciferase were mixed with 300 μl of Firefly Luciferase Assay Buffer. 25 μl of the mixture was then added to each 10 μl sample. Luciferase levels were immediately quantified using the luminescence setting of a Turner TD 20e luminometer. Twelve microlitres of Renilla were then mixed with 600 μl of Renilla Buffer and 25 μl of the mixture was added to the same wells. Renilla levels were detected using the same instrument. Sets of 6 or 9 wells were run at one time to minimize the time between substrate addition and detection of luminescence.

2.15 Avian retrovirus production

Infectious RCASBP retroviral particles were produced with the modified protocol of Gordon et al. (Gordon et al., 2010). Chicken DF-1 fibroblasts were seeded at 1 x 10⁶ cells per 10 cm dish and transfected the following day with a 10 μg RCAS plasmid using PEI. When cells reached confluence, the media was collected, the cells trypsinized, pooled with un-transfected DF-1 cells then re-seeded in three 15 cm dishes in the following proportions per 15 cm dish: one-third of a dish of transfected cells, one-third of the media from a dish of transfected cells, and one half of a sub-confluent dish of un-transfected cells. Twenty-four hours post-confluence, usually after 2-3 days, the media
<table>
<thead>
<tr>
<th>VISTA Enh. ID</th>
<th>Galgal6 coordinates (GRCg6a/galGal6)</th>
<th>Amplicon size (bp)</th>
<th>Primers (5’→ 3’)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>n/a</td>
<td>chr5:45912604-45913320</td>
<td>632</td>
<td>For:TATAGGTACCCTAAAGTTTTCATTCATATAA Rev:TATACTCGAGCCTTGCAAGTCAGGTCCTGAA</td>
<td>This study</td>
</tr>
<tr>
<td>n/a</td>
<td>chr4: 43390349-43390552</td>
<td>210</td>
<td>For:TATAGGTACCCTTGATATTTCTCTTCGCGC G Rev:TATACTCGAGCGCGTATCTCGGTGTAATTG</td>
<td>(Charité et al., 2001)</td>
</tr>
<tr>
<td>mm384</td>
<td>chr5:58582186-58582795</td>
<td>605</td>
<td>For:TATA GGTACC GCT GAT TAG TTA ATT AAC AGC C Rev: TATACTCGAGGCTGCTTGGC TGGAGTTG</td>
<td>Vista/face base hub</td>
</tr>
<tr>
<td>mm901</td>
<td>chr7:6244053-6245526</td>
<td>1468</td>
<td>For: TATA GGTACC CTT CCT GCC CTG ATT TAG GAC Rev: TATACTCGAGTCCATAAA ATT GGA CCA CA</td>
<td>Vista/face base hub</td>
</tr>
<tr>
<td>mm924</td>
<td>chr3:109433971-109434912</td>
<td>941</td>
<td>For:TATAGGTACCCTCCACTCGATGCTAGCAACT Rev:TATACTCGAGCGAGTTGGCAAATGCTTCTT</td>
<td>Vista / face base hub</td>
</tr>
<tr>
<td>mm1090</td>
<td>chr21:1351394-1352131</td>
<td>782</td>
<td>For: TATA GGTACC CTCTGCTGGCTGATGCTAAC Rev: TATACTCGAGAACACTCCTTGGGCTTCTCA</td>
<td>Vista / face base hub</td>
</tr>
<tr>
<td>mm1114</td>
<td>chr1:154020182-154020936</td>
<td>754</td>
<td>For: TATA GGTACC GCATGGCTCTAAAGCACTTTTT Rev:TATACTCGAGGATTAAGCAATGTTCATGTGA</td>
<td>Vista / face base hub</td>
</tr>
</tbody>
</table>
was changed to a low serum formulation (DMEM plus 1% FBS, 2 mM L-glutamine, 100U/mL penicillin, and 100 µg/ml of streptomycin). This media was collected and replaced with fresh low serum media 24 hours later for three consecutive days. Virus-containing media was stored at -80°C until concentrated, whereupon it was thawed and pre-cleared by centrifugation at 5,000 x g for 10 minutes at 4°C. A glass fibre pre-filter was added to a 0.45 µm Stericup filter, and the supernatant from all three collections was pooled and filtered. The filtered supernatant was divided into Ultra filter tubes (Beckman) and centrifuged in a SW-32 rotor at 32,000 x g at 4°C for 3 hours. The supernatant was decanted, and tubes were placed on an orbital shaker, in ice, for 2 hours at 150-200 rpm to facilitate resuspension of the virus pellet in the residual supernatant. Virus resuspension was completed by pipetting, pooled from all centrifuge tubes, and snap-frozen in single-use (10 µl) aliquots. Exogenous protein expression was confirmed by immunoblotting of protein extracts from terminal DF-1 cultures (see below).

2.15.1 Virus validation by western blotting

After the final virus collection, the cells from one 15 cm dish were washed in PBS and then scraped into 1ml of pre-chilled PBS. Cells were pelleted by centrifugation at 6,000 x g for 5 minutes at 4°C, then lysed in 1x Laemalli buffer (10% glycerol, 2% SDS, 5% β-mercaptoethanol, 0.05 M Tris-HCl pH 6.8) containing protease inhibitors (Complete Mini protease inhibitor, Roche). This cell lysate was sonicated for 3x 15 seconds, with an interval of 30 seconds, then heated for 5 minutes at 95°C. Protein concentration was determined using a Bicinchoninic acid (BCA) protein assay kit (Pierce). Forty micrograms of each protein sample were separated on a 13% SDS polyacrylamide gel at 200 V and then transferred onto a pre-soaked (methanol) PVDF membrane for 45 minutes at 300 mA. The membrane was rinsed 2x 5 minutes with PBT (PBS containing 0.1% Tween-20) and blocked with PBT plus 5% skim milk powder for 1 hour. Epitope-tagged virus-encoded proteins were detected with 9E10 α-Myc monoclonal antibody (hybridoma supernatant diluted 1:1,000 in PBT) overnight at 4°C.

The next day, membranes were washed 3x 10 minutes with PBT and incubated for one hour with goat anti-mouse HRP-conjugated secondary antibody (Invitrogen Cat #
Membranes were re-washed 3x 10 minutes and imaged with enhanced chemiluminescence (ECL) in Doc XRS using Image Lab software (BioRad). The same membrane was re-probed for actin to confirm equal protein loading. The membrane was washed 5x 5 minutes each and incubated for 30 minutes in stripping buffer (1M Tris-HCl pH 6.8, 20% SDS, 0.8% β-mercaptoethanol) at 50°C. Membranes were then washed 5x 10 minutes each in PBT and blocked with PBT containing 5% skim milk powder for 1 hour with continuous rocking. The primary β-actin antibody (Sigma Cat # 45441) was diluted in 1:1,000 in PBT containing 5% skim milk powder and incubated overnight at room temperature with continuous rocking. The next day, the membrane was washed 3x 10 minutes each with PBT and incubated for 1 hour with a secondary antibody; the membrane was then re-imaged as above.

2.15.2 Virus titration

DF-1 cells were seeded at 8 x 10⁴ cells/well in 12-well plates. Serial dilutions of the concentrated virus were prepared between 10⁻⁵ and 10⁻⁷, and 100 µl diluted virus was added to duplicate wells. Forty-eight hours later, cells were washed with PBS, fixed for 15 minutes with 4% PFA in PBS, re-washed with PBS, and blocked with PBST (PBS containing 10% chicken serum and 0.1% TritonX-100) for 10 minutes at room temperature. Cells were incubated with the 3C2 monoclonal antibody (hybridoma supernatuted 1:5 in PBT) for 30 minutes to detect the viral coat protein. Cells were rinsed thoroughly with PBST and incubated with a biotinylated anti-mouse IgG secondary antibody. Cells were re-washed with PBT (PBS containing 0.1% v/v Tween-20) and incubated with detection solution (Vectastain Elite ABC kit, Vector Laboratories) for 30 minutes. Cells were rinsed with PBT and incubated with DAB substrate using a DAB substrate kit (Vector Laboratories). When the desired signal-to-noise ratio was reached, cells were rinsed extensively with PBT. Virus-containing cell clones were counted and averaged. Virus preparations that fell in the range of 3 x 10⁸ to 2 x 10⁹ infectious units/ml were used for in ovo infection.
2.16 In Ovo infection

SPF eggs (Charles River, Ottawa) were incubated to HH16-17. The high-titer virus was mixed with 0.1 vol 1% Fast Green dye and injected into the mandibular and maxillary branches of the PA1 on the right side of the embryo. Eggs were sealed and incubated for an additional 48 hours for in situ treatment and harvested at E10 for skull staining.

2.17 Statistical analysis

All experiments were done a minimum of three times. Each qPCR data point represents the average of two duplicate assays. Each luciferase transcription data point represents the average of three technical replicates. The significance of the difference between the two experimental groups was determined with an unpaired two-tailed Student t-test. P < 0.05 was considered statistically significant.

2.18 Chemical and reagents

All reagents with their suppliers are tabulated below.

Table 2.6: Chemicals/ reagents and their supplier's list

<table>
<thead>
<tr>
<th>Reagent Name</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>3C2</td>
<td>Hybridoma</td>
</tr>
<tr>
<td>9E10</td>
<td>Hybridoma</td>
</tr>
<tr>
<td>Acetone</td>
<td>ThermoScientific</td>
</tr>
<tr>
<td>Agar</td>
<td>Fisher Bioreagents</td>
</tr>
<tr>
<td>Agarose</td>
<td>Ultra-Pure</td>
</tr>
<tr>
<td>Alcian Blue</td>
<td>Sigma</td>
</tr>
<tr>
<td>Alizarin Red</td>
<td>Sigma</td>
</tr>
<tr>
<td>BCA Kit</td>
<td>Pierce</td>
</tr>
<tr>
<td>BCIP</td>
<td>Roche</td>
</tr>
<tr>
<td>Bosentan</td>
<td>Toronto Research Chemicals</td>
</tr>
<tr>
<td>Corn Oil</td>
<td>Sigma</td>
</tr>
<tr>
<td>DAB Kit</td>
<td>Vector Labs</td>
</tr>
<tr>
<td>DIG RNA labelling Mix</td>
<td>Roche</td>
</tr>
<tr>
<td>DIG-AP-Fab fragment antibody</td>
<td>Roche</td>
</tr>
<tr>
<td>DMEM</td>
<td>ThermoScientific</td>
</tr>
<tr>
<td>DNA restriction enzymes</td>
<td>ThermoScientific</td>
</tr>
<tr>
<td>Item</td>
<td>Supplier</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>dNTP’s</td>
<td>Thermo Fisher</td>
</tr>
<tr>
<td>Fast Green</td>
<td>Fisher Bioreagents</td>
</tr>
<tr>
<td>FBS</td>
<td>Wisent</td>
</tr>
<tr>
<td>Formamide</td>
<td>ThermoScientific</td>
</tr>
<tr>
<td>Gene Ruler</td>
<td>ThermoScientific</td>
</tr>
<tr>
<td>Glutaraldehyde</td>
<td>Sigma</td>
</tr>
<tr>
<td>Glycerol</td>
<td>Fisher Chemicals</td>
</tr>
<tr>
<td>Glycine</td>
<td>Fisher Chemicals</td>
</tr>
<tr>
<td>Heparin</td>
<td>Sigma</td>
</tr>
<tr>
<td>Hydrogen peroxide</td>
<td>Fisher Chemicals</td>
</tr>
<tr>
<td>IgG Biotinlayted</td>
<td>Vector Labs</td>
</tr>
<tr>
<td>KCl</td>
<td>Fisher Chemicals</td>
</tr>
<tr>
<td>L- glutamine</td>
<td>Corning</td>
</tr>
<tr>
<td>Levamisole</td>
<td>Sigma</td>
</tr>
<tr>
<td>Maxi kit</td>
<td>ThermoScientific</td>
</tr>
<tr>
<td>Methanol</td>
<td>Fisher Chemicals</td>
</tr>
<tr>
<td>MgCl2</td>
<td>Fisher Chemicals</td>
</tr>
<tr>
<td>Midi Kit</td>
<td>ThermoScientific</td>
</tr>
<tr>
<td>NaCl</td>
<td>Fisher Chemicals</td>
</tr>
<tr>
<td>NBT</td>
<td>Roche</td>
</tr>
<tr>
<td>Pen/ Strep</td>
<td>Corning</td>
</tr>
<tr>
<td>PFA</td>
<td>Acros</td>
</tr>
<tr>
<td>Phenol: Chloroform: Isoamyl</td>
<td>Fisher Chemicals</td>
</tr>
<tr>
<td>Phusion Enzyme</td>
<td>ThermoScientific</td>
</tr>
<tr>
<td>Protease Inhibitor</td>
<td>Roche</td>
</tr>
<tr>
<td>Proteinase K</td>
<td>Roche</td>
</tr>
<tr>
<td>Qscript Mix</td>
<td>Quanta Bio</td>
</tr>
<tr>
<td>RapidOut DNA Removal Kit</td>
<td>Thermo Fisher</td>
</tr>
<tr>
<td>Red Safe</td>
<td>Frogga Bio</td>
</tr>
<tr>
<td>RNA Extraction kit</td>
<td>Qiagen</td>
</tr>
<tr>
<td>rRNasin</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>SDS</td>
<td>Fisher Bioreagents</td>
</tr>
<tr>
<td>SP6 RNA polymerase</td>
<td>ThermoScientific</td>
</tr>
<tr>
<td>Strata Clean</td>
<td>Agilent Tech.</td>
</tr>
<tr>
<td>SYBR</td>
<td>BioRad</td>
</tr>
<tr>
<td>T4 Ligase</td>
<td>ThermoScientific</td>
</tr>
<tr>
<td>T7 RNA polymerase</td>
<td>ThermoScientific</td>
</tr>
<tr>
<td>TOPO kit</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>Triton-X100</td>
<td>Fisher Bioreagents</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Trypsin</td>
<td>ThermoScientific</td>
</tr>
<tr>
<td>Tryptone</td>
<td>Fisher Bioreagents</td>
</tr>
<tr>
<td>TURBO Rnase/Dnase</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>Tween-20</td>
<td>Fisher Bioreagents</td>
</tr>
<tr>
<td>Vectastain Kit</td>
<td>Vector Labs</td>
</tr>
</tbody>
</table>
Chapter 3

*Dlx* Gene Expression in the Developing Chick Pharyngeal Arches and Dependance on Endothelin Signalling

Introduction

The vertebrate skull is a complex derivative of NCCs (Le Douarin et al., 2007) that contribute several bones and cartilages to the skull. NCC-derived craniofacial primordia arise, in part, from the interpretation of ventral endothelin signalling.

The endothelin signalling pathway comprises Edn ligands and their receptors, Ednra and Ednrb. Edn1 was discovered in 1988 (Yanagisawa et al., 1988) and is a 21 amino acid ligand derived from a 200 amino acid precursor protein (Clouthier et al., 1998, 2010) that undergoes three proteolytic treatments. The first cleavage is done by signal peptidase, which breaks the pre-proendothelin peptide into pre-endothelin. This pre-endothelin peptide further undergoes treatment by a furin peptidase to form a 38 amino acid inactive peptide. A metalloprotease, either Ece-1 or Ece-2, activates the peptide into the 21 amino acid, mature, Edn1 (Gaiano et al., 1999; Nakano et al., 1997). In the chick, there are three endothelin genes: *Edn1, Edn2*, and *Edn3* and three receptors (*EDNRA, EDNRB*, and *EDNRB2*), whereas mammals lack *EDNRB2* (Braasch & Schartl, 2014; Kempf et al., 1998; Lecoin et al., 1998; Liu et al., 2019). Mice deficient in Edn1 or Edn1-converting enzymes show malformations of craniofacial structures and homeotic transformation of the lower jaw into an upper jaw (Clouthier et al., 1998; Kempf et al., 1998; Ozeki et al., 2004; Ruest et al., 2004). The endothelin receptors are expressed in migrating CNCC and in the ectomesenchyme of PA1, whereas the endothelin signalling molecule is specifically secreted by the epithelium and mesoderm of md-PA1 (Clouthier et al., 1998).
Liu and colleagues showed that Ednr formed heterodimers with $G_\alpha$ and $G_\beta$ proteins and stimulated phospholipase C, which is capable of activating several signalling pathways, including calcium transportation, MAPK/ERK, and cAMP/PKA in mammals and birds (Liu et al., 2019). Ednra signalling is therefore essential in eliciting the expression of several TFs in PA1 in a timely manner (Cesario et al., 2015; Clouthier et al., 1998; Ozeki et al., 2004; Ruest et al., 2004; Thomas et al., 1998).

In mice, Ednra is required for arch patterning between E8.5 and E9.0 for a window of approximately 18 hours (Ruest & Clouthier, 2009) and acts through the activation of various lower jaw-specific genes, mostly TFs, such as $Dlx5$, $Dlx6$, $Gsc$, $Pitx1$ and $Hand2$ (Clouthier et al., 1998; Ozeki et al., 2004; Ruest et al., 2004). Among them, $Dlx5$ and $Dlx6$ play a crucial role in arch patterning and jaw specification. Conversely, ectopic expression of Edn1 in the upper jaw transforms the upper jaw into a lower jaw (Sato et al., 2008) and activates its transcriptional targets, such as $Dlx5$, $Dlx6$, $Gsc$, and $Pitx1$. A series of experiments were done in mouse embryos treated with the pharmacological inhibitor of Ednra, Bosentan, at different developmental time points. Gene expression analysis of some Ednra-dependent transcription factors, such as $Dlx5$, $Dlx6$, and $Hand2$, showed downregulation with craniofacial defects around E8.5 – E9 when treated with Bosentan. Comparatively, treatment after E9 showed neither defects in craniofacial structures nor altered expression of any Edn-dependent genes (Ruest et al., 2005). These findings indicate that Dlx5 and Dlx6 carry out Endothelin-independent functions and that Edn signalling is not required for the maintenance of $Dlx$ gene expression in the jaw-forming tissue. It is also interesting to note that, while $Endra$ and $Ednrb$ are both expressed in the PA1, they do not show reciprocal functional redundancy (Hosoda et al., 1994). Unlike $Ednra$ mutant mice, $Ednrb$ mutants do not display any craniofacial malformation (Hosoda et al., 1994). While these findings could imply that $Ednra$ is compensating for the absence of $Ednrb$, the reciprocal compensation does not occur (Clouthier et al., 2010; Ruest & Clouthier, 2009). Thus, $Ednrb$ is not involved in mandibular arch patterning, and endothelin-responsive genes depend entirely on $Ednra$ (Clouthier et al., 2010; Kurihara et al., 1994).
*Hox* genes are involved in anterior-posterior patterning of the main body axis (Merlo et al., 2000; Zappavigna et al., 1994) and are expressed in neural crest cells (Parker et al., 2018), the cranial neural crest cells that migrate into PA1 are *Hox*-negative (Gendron-Maguire et al., 1993; Hunt et al., 1991; Rijli et al., 1993; Rivera-pérez et al., 1995). Instead, another homeobox gene family, *Dlx*, patterns the jaw-forming tissue (Merlo et al., 2000).

Hu et al. (2015) found an endothelin-responsive myocyte enhancer factor 2 (Mef2C) enhancer in the third intron of the *Mef2C* gene. This enhancer lost activity when embryos were treated with Bosentan, suggesting a direct dependence of this enhancer on endothelin in post-migratory CNCCs (Hu et al., 2015). Edn signalling and Mef2C in mice and zebrafish are involved in *Dlx* gene expression (Hu et al., 2015; Miller et al., 2007). Furthermore, *eden1* mutants in zebrafish show an extreme phenotype in dorso-ventral jaw patterning. These mutants lack jaws and have suckers, a characteristic of agnathans (Miller et al., 2000).

The “hinge and caps” model of jaw structure explains jaw patterning and the functional registration of the jaw tissues. This model refers to a “hinge’ which is a common factor of the upper (maxilla) and lower (mandibles) jaw. This junction of the maxillary and mandibular branches allows for the articulation of the jaws, whereas, the caps are the jaw itself (Compagnucci et al., 2013; Michael J Depew et al., 2005). The functional integration of both components is the basis of the modern gnathostome's jaw articulation. The hinge location is determined by Fgf8 signalling, whereas *Dlx* genes govern cap patterning. *Dlx* genes in PA are regulated by endothelin signalling. In vertebrates, six evolutionary conserved *Dlx* genes have been identified (*Dlx1-Dlx6*). In mice, experimentation has shown that *Dlx5* and *Dlx6* are downstream of Edn signalling (Clouthier et al., 1998; Ozeki et al., 2004; Ruest et al., 2004), as both *Dlx5/6* and *Ednra* null embryos share the same mandibular-maxillary transformed phenotype (Beverdam et al., 2002; Depew et al., 2002; Ozeki et al., 2004; Ruest et al., 2004).

Contrary to mammals and birds, zebrafish have eight *dlx* genes (Verreijdt et al., 2006). The additional genes arose from a genome duplication in teleosts (Ekker et al., 1992
1994; Ellies, Stock, et al., 1997). As for their mammalian orthologues, six are expressed in convergently transcribing bigene clusters, where \textit{dlx}1\textit{a} is linked with \textit{dlx}2\textit{a}, \textit{dlx}3\textit{b} with \textit{dlx}4\textit{b}, and \textit{dlx}5\textit{a} with \textit{dlx}6\textit{a}. The two remaining genes, \textit{dlx}2\textit{b} and \textit{dlx}4\textit{a}, do not seem to be linked to other \textit{dlx} genes (Ellies, et al., 1997).

In zebrafish embryo at 36 hours post-fertilization (hpf), \textit{dlx}2\textit{a} is expressed along the dorsoventral axis of [upper and lower PA1 primordia] pharyngeal arches, overlapping with the expression of \textit{dlx}1\textit{a} but more strongly expressed than \textit{dlx}1\textit{a} (Coffin-Talbot et al., 2010). The other linked pair \textit{dlx}3\textit{b}-\textit{dlx}4\textit{b} is expressed in the intermediate zone at the same developmental timepoint. \textit{Dlx}5\textit{a} and \textit{dlx}6\textit{a} are also expressed in the dorsal part of the PA but, like \textit{dlx}1\textit{a}, \textit{dlx}6\textit{a} shows weaker expression and is more restricted than \textit{dlx}5\textit{a}. Zebrafish \textit{dlx} genes also show nested expression. Talbot et al (2010) showed the expression of \textit{dlx}3\textit{b}, \textit{dlx}4\textit{b}, and \textit{dlx}5\textit{a} in the intermediate zone. The unlinked \textit{dlx} gene, \textit{dlx}4\textit{a}, is restricted to the intermediate zone whereas \textit{dlx}2\textit{b} is not expressed in the arches at all (Talbot et al., 2010).

At E10.5, in mice, all six genes are expressed in PAs. \textit{Dlx}1 and \textit{Dlx}2 are expressed more distally in PA1, in both maxillary and mandibular parts, and PA2. \textit{Dlx}5 and \textit{Dlx}6 are also expressed in an overlapping fashion at this developmental time point but the expression is restricted to the md-PA1. \textit{Dlx}3 and \textit{Dlx}4 expression is also restricted to md-PA1 and only in more distal tissue (Jeong et al., 2008).

Expression of the \textit{Dlx} gene family has been well described in mouse and zebrafish embryos, and numerous studies have documented the functions of these genes in jaw patterning. Fewer genetic studies have been conducted on the chick embryo and \textit{Dlx} gene expression has not been systematically studied in this important model organism. Comparative embryology among model and non-model organisms offers a way to identify conserved features of gene expression, as well as lineage-specific differences. In this study, the tempo-spatial expression of the \textit{Dlx} gene family was elucidated at various developmental time points using WMISH. Specifically, we describe and compare \textit{Dlx} gene expression patterns for all six family members in PA1 after
significant NCC migration has ended while jaw patterning is active. Secondly, the
dependence of \textit{Dlx} gene expression on the endothelin signalling pathway was assessed
for each family member by blocking the Edn pathway using Bosentan, an antagonistic
inhibitor of Ednra. Gene expression in control and Bosentan-treated chick embryos was
then compared by WMISH and qPCR. In the chick, the status of \textit{Dlx4} is enigmatic. A
complete cDNA has not been cloned (Brown et al., 2005; Takechi et al., 2013) and no
expression has been described. Here, we use an exon 3-specific probe, amplified from
genomic DNA, to provide the first evidence of \textit{Dlx4} expression in the chick embryo.

3.1 Results

Before describing the Spatio-temporal expression of \textit{Dlx} genes in chickens, it is essential
to explain a few terms used when describing expression patterns in the PA (Fig 3.1).

3.1.1 \textit{Dlx1} & \textit{Dlx2} expression in chick pharyngeal arches

The overall \textit{Dlx1} expression is summarized in Fig 3.2. At HH16 (Fig 3.3 A-D), \textit{Dlx1} (Fig
3.3 A- T) is expressed in PA1-PA4, highest in PA1, and progressively lower expression
in more caudal arches, being only weakly detected in PA4. In each arch, expression is
highest at the mesenchymal margins of the arch, being reduced or absent in the
mesodermal core, and there is no expression in the surrounding ectoderm (Fig 3.3 A &
B). Viewed ventrally, expression in the mandibular arch extends in a thin arc that extends
the entire length of the oral margin to the distal aboral half (Fig 3.3 C & D). Expression in
the maxillary branch is more uniform throughout, highest in the distal half.
Expression is proportionally higher in all tissues at HH18 (Fig 3.3 E-H), such that expression can now be detected in PA5 but shows the same patterns (Fig 3.3 F-H). Expression continues to increase throughout PA1, such that by HH20 (Fig 3.3 I-L) high levels of Dlx1 transcripts are found throughout the maxillary branch to the lambdoid
junction and in most of the mandibular branch, with the exception of a distal-aboral zone (Fig 3.3 K & L). Expression at HH22 (Fig 3.3 M-P)) appears the same as at HH20 when viewed laterally (Fig 3.3 M), but the ventral regions of the first arch reveal a more complex pattern. Specifically, as the mandibular tissue elongates, higher-level expression is maintained in the oral half and proximally, at the hinge, with lower expression in an expanded distal-aboral zone and just distal to the hinge. Expression in the maxillary tissue now shows a more obvious demarcation between a proximal zone that extends distally along the oral margin and the rest of the maxillary arch, where expression is reduced compared to HH20.

The overall Dlx2 expression is summarized in Fig 3.4. Dlx2 expression (Fig 3.5 A-T) at HH16 (Fig 3.5 A-D) largely overlaps with Dlx1, although expression in PA2-5 is relatively higher such that all four arches express Dlx2 to a similar degree. Subsequently, Dlx2 expression is more restricted in PA1 compared with its cis-linked paralogue. Specifically, in the maxillary arch at HH18 (Fig 3.5 E-H), Dlx2 is expressed in a more restricted zone along the oral margin. Restriction of high-level Dlx2 expression to the oral mesenchyme continues through HH24 (Fig 3.5). In the mandibular branch, broader Dlx2 expression at HH18 becomes progressively more restricted to a distal-oral zone. A similar refinement of Dlx2 expression occurs in PA2; initial expression throughout the ectomesenchyme becomes restricted to both distal and caudal domains.
Fig. 3.2: *Dlx1* general expression summary

(A) Schematic demonstrating the overall expression (representative of most of the described stages) of *Dlx1* in PA1 (magenta) and PA2 (yellow). (B) describes the expression in caps and hinges at most of the developmental stages between HH16-HH26. (C) WMISH of chick embryo at HH20, whole lateral view of the embryo is shown. (D) Zoomed view of PA’s in chick embryo at HH20. (E) Frontal view of HH20 chick embryo. Al, allantois; di, diencephalon; ey, eye; fl, fore limb; h, heart; hl, hind limb; mes; mesencephalon; PA (I, II, III, IV, V), tel, telencephalon; md, mandible; mx, maxilla; ov, otic vesicles; so, somites; λ, lambdoidal junction (where maxilla meets fronto-nasal prominence).
Fig. 3.3: *Dlx1* expression in the chick embryo during pharyngeal arch patterning.

*Dlx1* expression patterns of the chick embryo in the whole body, branchial arches, nasal placode, and otic vesicles were viewed by WMISH. The Hamburger-Hamilton stage is indicated. All lateral images are of the right side. Each row represents different views of the same embryo.
Fig. 3.4: Dlx2 general expression Summary

(A) Schematic demonstrating the overall expression (representative of most of the described stages) of Dlx2 in PA1 (pink) and PA2 (yellow). (B) Describes the expression in caps and hinges at most of the developmental stages between HH16-HH26. (C) WMISH of chick embryo at HH20, whole lateral view of the embryo is shown. (D) Zoomed view of PA’s in chick embryo at HH20. (E) Frontal view of HH20 chick embryo. Al, allantois; di, diencephalon; ey, eye; fl, fore limb; h, heart; hl, hind limb; mes, mesencephalon; PA (I, II, III, IV, V), tel, telencephalon; md, mandibles; mx, maxilla; ov, otic vesicles; so, somites; λ, lambdoidal junction (where maxilla meets fronto-nasal prominence).

3.1.2 Dlx3 & Dlx4 expression in chick pharyngeal arches

The overall Dlx3 expression is summarized in Fig 3.6. The early post-migratory expression of Dlx3 (Fig 3.7 A-T) is most obvious in the first two pharyngeal arches at HH17 (Fig 3.7 A-D). Weaker expression in PA3 and PA4 at this time is further reduced at HH18 (Fig 3.7 E-H) and disappears by HH20 (Fig 3.7 I-L). In the first arch, early expression of Dlx3 resembles that of Dlx2, being expressed throughout the mandibular and maxillary ectomesenchyme and in the frontonasal ectoderm beyond the lambdoidal junction at HH17. Significant expression is also noted in the otic vesicle and in the surface ectoderm corresponding to the nasal placodes at HH17 and HH18.
Fig. 3.5: *Dlx2* expression patterns in the chick embryo

*Dlx2* expression patterns of the chick embryo in the whole body, branchial arches, nasal placode, and otic vesicles were viewed by WMISH. The Hamburger-Hamilton stage is indicated. All lateral images are of the right side. Each row represents different views of the same embryo.
Viewed ventrally, *Dlx3* expression in mdPA1 resolves to a discrete aboral domain midway along the proximal-distal axis. Expression in PA2 is similarly restricted proximal-distally. Expression of *Dlx3* in mx-PA1 is no longer detectable by HH20 while expression of *Dlx3* continues in mdPA1, strongest in the medial-aboral zone that expands as the mandibular tissue elongates. By HH26 (Fig 3.7 Q-T) there is, additionally, a novel expression in the distal oral ectoderm of mdPA1. Expression in the otic vesicle is down-regulated between HH22 and HH26 but remains high in the nasal pits (Fig 3.7 S).

**Fig. 3.6: Dlx3 general expression Summary**

(A) Schematic demonstrating the overall expression (representative of most of the described stages) of *Dlx3* in PA1 (mauve) and PA2 (yellow). (B) describes the expression in caps and hinges at most of the developmental stages between HH16-HH26. (C) WMISH of chick embryo at HH20, whole lateral view of the embryo is shown. (D) Zoomed view of PA’s in chick embryo at HH20. (E) Frontal view of HH20 chick embryo. Al, allantois; di, diencephalon; ey, eye; fl, fore limb; h, heart; hl, hind limb; mes, mesencephalon; PA (I, II, III, IV, V), tel, telencephalon; md, mandibles; mx, maxilla; np, nasal prominence; ov, otic vesicles; so, somites; λ, lambdoidal junction (where maxilla meets fronto-nasal prominence).
Fig. 3.7: *Dlx3* expression patterns in the chick embryo

*Dlx3* expression patterns of the chick embryo in the whole body, branchial arches, nasal placode, and otic vesicles were viewed by WMISH. The Hamburger-Hamilton stage is indicated. All lateral images are of the right side. Each row represents different views of the same embryo.
The overall *Dlx4* expression is summarized in Fig 3.8. *Dlx4* (Fig 3.9 A-X) expression was not detected until HH19 (Fig 3.9 A-D) when transcripts were weakly detected in a small aboral domain of mdPA1 and a rostral domain within PA2 (Fig 3.9 C-D). Viewed ventrally, both *Dlx4* expression domains were distinctly smaller and fit within the zones of *Dlx3* expression in PA1 and PA2. Expression of *Dlx4* continues in these discrete zones in mdPA1 and PA2 until HH24, but the expression is not seen elsewhere in the pharyngeal or frontonasal tissues (Fig 3.9 U-X).

**Fig. 3.8: Dlx4 general expression summary**

(A) Schematic demonstrating the overall expression (representative of most of the described stages) of *Dlx4* in PA1 (blue) and PA2 (yellow). (B) describes the expression in caps and hinges at most of the developmental stages between HH16-HH26. (C) WMISH of chick embryo at HH20, whole lateral view of the embryo is shown. (D) Zoomed view of PA’s in chick embryo at HH20. (E) Frontal view of HH20 chick embryo. Al, allantois; di, diencephalon; ey, eye; fl, fore limb; h, heart; hl, hind limb; mes; mesencephalon; PA (I, II, III, IV, V), tel, telencephalon; md, mandibles; mx, maxilla; np, nasal prominence; ov, otic vesicles; so, somites; λ, lambdoidal junction (where maxilla meets fronto-nasal prominence).
Fig. 3.9: *Dlx4* expression patterns in the chick embryo

*Dlx4* expression patterns of the chick embryo in the whole body, branchial arches, nasal placode, and otic vesicles were viewed by WMISH. The Hamburger-Hamilton stage is indicated. All lateral images are of the right side. Each row represents different views of the same embryo.
3.1.3 *Dlx5 & Dlx6* expression in chick pharyngeal arches

The overall *Dlx5* expression is summarized in Fig 3.10. Like *Dlx1* and *Dlx2* expression at HH16, *Dlx5* (Fig 3.11 A-T) is expressed in the ectomesenchyme of PA1-4 and, like *Dlx1*, *Dlx5* expression (Fig 3.11 A-D) is highest in PA1, albeit restricted to the mandibular branch. While *Dlx5* expression extends into proximal regions of PA1, expression is lower in the proximal ectomesenchyme of PA2-4. In all arches, *Dlx5* expression is low in the most distal regions (Fig 3.11 A-T). Additionally, *Dlx5* is expressed in the otic vesicle and nasal pits. This pattern of expression continues through HH18 (Fig 3.11 E-H) and HH20 (Fig 3.11 I-L) in the pharyngeal tissue, maintained at high levels throughout the first two arches and becoming restricted rostrally in PA3 and PA4. By HH22 (Fig 3.11 M-P), *Dlx5* was upregulated in most oral ectomesenchyme of the maxillary branch and expression continues to increase until, at HH26 (Fig 3.11 Q-T), expression along the oral margin of the maxillary tissue is as robust as the mandibular expression. Most of the developing mandible expresses *Dlx5* at a high level, except for the most distal margin.

**Fig. 3.10: Dlx5 general expression summary**

(A) Schematic demonstrating the overall expression (representative of most of the described stages) of *Dlx5* in PA1 (orange) and PA2 (yellow). (B) describes the expression in caps and hinges at most of the developmental stages between HH16-HH26. (C) WMISH of chick embryo at HH20, whole lateral view of the embryo is shown. (D) Zoomed view of PA’s in chick embryo at HH20. (E) Frontal view of HH20 chick embryo. Al, allantois; di, diencephalon; ey, eye; fl, fore limb; h, heart; hl, hind limb; mes; mesencephalon; PA (I, II, III, IV, V); tel, telencephalon; md, mandibles; mx, maxilla; ov, otic vesicles; so, somites; λ, lambdoidal junction (where maxilla meets fronto-nasal prominence).
Fig. 3.11: *Dlx5* expression patterns in the chick embryo

*Dlx5* expression patterns of the chick embryo in the whole body, branchial arches, nasal placode, and otic vesicles by WMISH. The Hamburger-Hamilton stage is indicated. All lateral images are of the right side. Each row represents different views of the same embryo.
The overall $Dlx6$ expression is summarized in Fig 3.12. $Dlx6$ is essentially co-expressed with $Dlx5$ in the pharyngeal ectomesenchyme and placodal ectoderm with the exception of the maxillary arch, in which expression of $Dlx6$ is restricted to a small proximal domain (Fig 3.11 A-T).

![Fig. 3.12: $Dlx6$ general expression summary](image)

(A) Schematic demonstrating the overall expression (representative of most of the described stages) of $Dlx6$ in PA1 (red) and PA2 (yellow). (B) describes the expression in caps and hinges at most of the developmental stages between HH16-HH26. (C) WMISH of chick embryo at HH20, whole lateral view of the embryo is shown. (D) Zoomed view of PA’s in chick embryo at HH20. (E) Frontal view of HH20 chick embryo. Al, allantois; di, diencephalon; ey, eye; fl, fore limb; h, heart; hl, hind limb; mes; mesencephalon; PA (I, II, III, IV, V), tel, telencephalon; md, mandibles; mx, maxilla; ov, otic vesicles; so, somites; \(\lambda\), lambdoidal junction (where maxilla meets fronto-nasal prominence).

3.1.4 Endothelin-responsiveness of $Dlx$ gene expression in the pharyngeal arches

$Dlx5$, $Dlx6$ and $Dlx3$ expression has previously been shown to be dependent on endothelin signalling in md-PA1 of mouse embryos (Ozeki et al., 2004; Ruest et al., 2004).
Fig. 3.13: *Dlx6* expression patterns in the chick embryo

*Dlx6* expression patterns of the chick embryo in the whole body, branchial arches, nasal placode, and otic vesicles by WMISH. The Hamburger-Hamilton stage is indicated. All lateral images are of the right side. Each row represents different views of the same embryo.
I treated chick embryos at HH13 with Bosentan, a pharmacological antagonist of Ednra (Kempf et al., 1998). Embryos were collected 24 hr post-treatment, (~HH17) for in situ analysis, 48 hr post-treatment (HH21-22) for qPCR, and 9 days post-treatment (E12) for skeletal staining (Fig. 3.14).

Fig. 3.14: Bosentan-treated jaw phenotype

(A) Control (vehicle-treated) chicken embryo at E12; yellow arrow head points to the lower beak. (B) hypoplasia of lower jaw (yellow arrowhead) upon Bosentan treatment, an inhibitor of endothelin signalling (C) Control (left) and Bosentan-treated (right) chicken mandibular jaw staining showing Alcian blue-stained cartilage and Alizarin red-stained bone. Note the duplication of the pterygoid in the treated mandible. Duplication is shown with *. mc, Meckel’s cartilage; pt, pterygoid; q, quadrate.

In agreement with earlier studies in chick embryos (Kempf et al., 1998), the expression of Dlx5 and Dlx6 were both reduced in Bosentan-treated embryos compared to vehicle-treated controls (Fig. 3.15, compare Q & R with S & T and U & V with W & X respectively) thereby verifying the dosage and timing of drug delivery. When the expression of Dlx3 was examined by in situ hybridization (Fig. 3.15, I-L), it was found to be nearly undetectable in the pharyngeal arches of Bosentan-treated embryos (Fig. 3.15, compare I & J with K & L). Dlx3 expression was less perturbed in the otic vesicles of Bosentan-treated embryos. Dlx4 expression was difficult to detect before HH19 so it was not possible to conclude anything about the endothelin dependence of Dlx4 using HH17 embryos (Fig. 3.15, M-P). To complete the analysis, I also examined the expression of Dlx1 and Dlx2 following Bosentan treatment. Interestingly, blocking endothelin signalling did not result in reduced expression of either gene.
**Fig. 3.15: Dlx gene response to reduced endothelin signalling**

Differential expression of putative Dlx genes at 24 hr post-Bosentan treatment. Embryos were treated as described. Dlx1 control (A & B) and Dlx1 Bosentan treated (C & D) 3 of 4 embryos showed downregulation. Dlx2 (E & F control) and (G & H treated) 4 of 4 embryos showed downregulation of Dlx2. Dlx3 (I & J control) whereas (K & L treated) in total 4 of 4 embryos showed downregulation. Dlx4 (M & N control) and (O & P Bosentan treated embryos) Dlx4 expresses beyond HH18. Dlx5 (Q & R control) and (S & T treated) 4 of 4 embryos exhibited downregulation of Dlx5. Dlx6 (U & V control) and (W & X treated) 3 of 3 embryos exhibited downregulation.
Fig. 3.16: Differential expression of $Dlx$ genes by qPCR analysis

Differential expression of $Dlx$ genes upon inhibition of endothelial signalling pathway by Bosentan. Gene expression levels in md-PA1 and PA2 at HH21-22, 48 hr post-treatment. The whole mandibular process and hyoid arch were excised from treated and untreated embryos and pooled from 2-3 embryos. The gene of interest was normalized to Succinate Dehydrogenase ($SDHA$) expression. Control gene expression is set at 1 for each gene. Control bars are shown in blue while orange bars are for bosentan treated embryos. Error bars denote the SEM. The experiment has been repeated a minimum of 3 times. *$P<0.05$, **$P<0.01$.

Rather, there appeared to be increased transcript abundance in the md-PA1 of Bosentan-treated embryos, as well as in PA2 and PA3 for both $Dlx1$ and $Dlx2$ (Fig. 3.15, compare A & B with C & D and E & F with G & H) respectively.

We next compared transcript levels by qPCR at a later stage of development. The mandibular and hyoid tissue was dissected and pooled from groups of 2-3 vehicle- or Bosentan-treated embryos, 48 hours post-treatment, at HH21-22. Although I could not find suitable qPCR primers for $Dlx2$, $Dlx1$ transcript abundance increased in the presence of Bosentan, to almost twice the level in control embryos (Fig. 3.16). All other $Dlx$ genes had reduced expression in Bosentan-treated embryos (Fig. 3.16). These data, therefore,
corroborated the response of individual Dlx genes to the down-regulation of endothelin signalling.

3.2 Discussion

Jaws are significant in gnathostome diversification, and Dlx genes are at the core of this innovation. In this study, I have demonstrated that, in the chick, all six Dlx genes are expressed in PA1. I also used the antagonist of endothelin signalling, Bosentan, and showed the endothelin dependence of each Dlx family member. Here, I compare the expression of chick Dlx genes with other taxa such as mouse and lesser spotted dogfish (Scyliorhinus canicula) shark.

3.2.1 Similarities and differences between paralogues in md-PA1

Dlx1 and Dlx2 show similar expression patterns in the whole body, specifically in the pharyngeal arches. While Dlx1 is expressed in the entire length of md-PA1, Dlx2 shows a more restricted expression distally in md-PA1 at HH20-22. Dlx3 is expressed in the branchial arches, sensory placodes, and limbs, whereas Dlx4 is only expressed in md-PA1 and the hyoid arch (Fig 3.7 & 3.9). The Dlx4 expression appears to be nested within the proximal-aboral domain of Dlx3 expression in md-PA1 and within the proximal-rostral domain in PA2. Dlx5 and Dlx6 expressions are the most similar among cis-linked paralogues. Dlx6 expression is weaker than Dlx5 in the sensory placodes at HH16-HH17.

Interestingly, At HH22, Dlx5 expression extends into mx-PA1 and by HH26 is expressed along the oral mesenchymal margin. In contrast, Dlx6 expression was not detected in mx-PA1 until HH26 where it was restricted to a small proximal domain (Fig 3.11 & 3.13). Also, the Dlx4 expression is nested within Dlx3, whereas Dlx2 and Dlx3 expressions are nested within Dlx1, Dlx5, and Dlx6. This overlapping expression is consistent with functional redundancy. In mice, contrary in chick, Dlx3 and Dlx4 expressions are nested with Dlx1 and Dlx2 in the mandibular branch where Dlx1 and Dlx2 expression is nested within a broader Dlx5 and Dlx6 expression domain.
3.2.2 Comparison with orthologues

Herein, I will compare chicken Dlx gene expression with sharks and mice to the extent that published data allows. In the shark I compared stages 20-27, focusing on stage 25, and mouse E10.5. Dlx1 and Dlx2 expression in chick (this study) and mouse (Jeong et al., 2008) PA1 are maxillo-mandibular. Shark Dlx1 at stages 22-25 is only expressed in the maxillary branch (Takechi et al., 2013). In the shark at stages 20 and 25, Dlx2 expression is maxilla-mandibular like in chicken and mice but interestingly at stage 27 expression is restricted to the maxillary arch (Compagnucci et al., 2013), but a weaker expression is detected in the mandibular branch at stage. In contrast, in the shark after stage 25, Dlx2 expression is more robust in the maxillary arch (Compagnucci et al., 2013). A low transcript level is detected in md-PA1. Chicken Dlx1 & Dlx2 show an overall hinge-cap-centric pattern that shows conservation with dogfish sharks and mice. Dlx3 expression is overall conserved among all three species in the mandible. Interestingly, Dlx3 has also been expressed in the shark (stage 21, 25, 27)(Compagnucci et al., 2013) and mice (J. Jeong et al., 2008) maxilla but not in the chick. Moreover, Dlx3 shows more transcript in md-PA1 in mice and shark taxa as compared to mx-PA1. In contrast to mice and chickens, in sharks, Dlx4 is expressed in PA1-PA4 and in the oropharyngeal cavity (Takechi et al., 2013). Chick Dlx3 and Dlx4 are cap-centric, as are in shark and mouse. Dlx5 and Dlx6 expression in chicken, mice, and sharks follow a conserved expression pattern and are expressed in the mandibular branch. Interestingly, Chicken HH22-25 and sharks stage 25 and 27 show expression of Dlx5 in the maxilla as well. Chicken Dlx5 and Dlx6 show a cap-hinge pattern seen in sharks and mice. All Dlx gene expression in hinges and caps is summarized in Fig. 3.17.

3.2.3 Endothelial maxillary-mandibular switch

We used Bosentan as a pharmacological inhibitor of endothelin signalling to monitor endothelin-dependent aspects of Dlx gene expression. Dlx5 and Dlx6 are conserved targets of endothelin signalling in mice and chicks. Here I observed that expression of Dlx1 and Dlx2 increased when endothelin signalling was antagonized. Conversely,
Overall the Dlx family is very well conserved in chicken at HH24-25, in the mouse at E10.5 and in shark at stage 25. The only exceptions are Dlx3 expression in mice and sharks in mx-PA1 and chicken Dlx5 expression in mx-PA1. The position of the hinges is shown in a circle whereas two caps, corresponding to the lower (mandiblar) and upper (maxillary) jaw are shown in elongated structures. Expression is shown with lavender shading.
Dlx3-Dlx6 were dependent on endothelin signalling for their expression in md-PA1 in the chick embryo. Specifically, Dlx3, Dlx5 and Dlx6 transcripts were undetectable by WMISH in md-PA1 of Bosentan-treated chicken embryos and all four Dlx genes (Dlx3-Dlx6) were down-regulated when measured by qPCR. Unlike Dlx5 and Dlx6, Dlx3 expression was not obviously changed in the otic vesicle of Bosentan-treated embryos, suggesting its regulation may be independent of endothelin signalling in that tissue.

In mice and zebrafish, Dlx expression is also endothelin-1 dependent (Clouthier et al., 2010; Miller et al., 2000). The expression of dlx3a is lost in the zebrafish sucker mutant, (Miller et al., 2000) and in the pharyngeal arches of Edn1 mouse mutants (C. T. Miller et al., 2000), whereas Dlx2 expression is disrupted in the hyoid arch of mice (Clouthier et al., 2000). Our Bosentan findings in chick embryos are in accordance with Bosentan findings in mice where Dlx5 is downregulated upon Bosentan treatment, but Dlx2 does not downregulate in mice spatially (Park et al., 2004). In short, in the absence of the endothelin-Dlx5/6, the mandibular jaw acquires a maxillary fate (Fig 3.14), which illustrates the significance of the endothelin signalling pathway and its conservation among vertebrates.

3.2.4 Dlx genes have nested expression patterns in chick PA1

As seen in other species, the chick Dlx gene family shows nested expression in the PA1. In md-PA1, Dlx5 and Dlx6 are expressed throughout the arch, whereas the remaining four genes of the family show expression within the domains of these two genes. This also strengthens the hypothesis that Dlx5 and Dlx6 regulate these Dlx genes in md-BA1.
Fig. 3.18: Schematic illustrating transformation of the lower jaw into an upper jaw

*Dlx* gene expression in chick PA1 at HH16-17. (A) A labelled head is shown (B) indicates the expression of five *Dlx* genes (*Dlx*4 did not express yet) in the PA1 of chick before endothelin inhibitor treatment. (C) Shows the loss of gene expression upon endothelin signalling pathway inhibition with Bosentan. The magenta color shows the expression of *Dlx*1 and *Dlx*2, red shows the expression of *Dlx*5 and *Dlx*6, whereas blue shows the expression of *Dlx*3. At HH17 *Dlx*4 is not expressed in the chick arches. The colored lines indicate the extent of expression along the proximal-distal axis of md-PA1.

However, misregulation of *Dlx*3 upon *Dlx*5/6 absence also suggests that *Dlx*5/6 is involved in *Dlx*3 regulation if it is not a potential direct target of endothelin signalling in PA1 at later stages.

### 3.2.5 Are members of a bigene pair co-expressed or not?

In this study, I observed that the expression of *cis*-linked *Dlx* paralogues was not always the same in the pharyngeal tissue, though they share one or more common intragenic enhancers. For example, *Dlx*1 and *Dlx*2 expression is similar but their patterns are not completely overlapping in PA1 or PA2. Similarly, *Dlx*4 does not show the same expression pattern as *Dlx*3 in these tissues. In contrast, *Dlx*5 and *Dlx*6 expression substantially overlap in md-PA1 (Fig 3.19 A), as well as in other tissues. Though the exact
mechanism by which one gene is expressed where the cis-linked paralogue is not, we can hypothesize that this situation is due to (1) one gene of the bi-gene cluster having a stronger and more sensitive promoter in some tissues (Fig 3.19 B) (2) having multiple enhancers or novel enhancer for activation other than an intragenic enhancer in a specific tissue (Fig 3.19 C) (3) promoter modification such as methylation of one paralogue’s promoter or the presence of a heterologous promoter (Fig 3.19 D). Model B is supported by the observation that the *Dlx3* expression domain is broader than that of *Dlx4* in md-PA1 (Fig. 3.7 and 3.9). Model C is supported by the observation that the 156i enhancer can recapitulate the full expression of *Dlx5* and *Dlx6* in the pharyngeal arches of transgenic mice while the 112a *Dlx1/2* intragenic enhancer can only partially recapitulate the *Dlx1* and *Dlx2* expression in the same tissue (Park et al., 2004). There is currently no evidence for model D. It can be concluded that bi-gene pairs are expressed in overlapping, but not identical, patterns.

**Fig. 3.19:** Schematic demonstrating possibilities of bigene cluster anomalous expression in tissue  
(A) Shows bigene clusters regular expression where both promoter and enhancer work together to express the target. (B) Despite the occupancy of intragenic enhancers, some promoters are sensitive or more strong  
(C) One of the enhancers in the bigene cluster is stronger (more acetylated) than the other. (D) Promoter modification such as methylation of histone and/or promoter. Created with BioRender.com.
3.3 Conclusion:

This study aimed to describe the expression of the Dlx gene family in the context of its pivotal role in PA1 morphogenesis and to test the response of each gene to reduced endothelial signalling. Here, for the first time, I showed the expression of Dlx4 in the chicken embryo. With some deviation from expression patterns in orthologues, and having some unique expression patterns in chicken, the overall Dlx gene expression profile is well conserved among species in terms of spatial-temporal expression and endothelin-dependence acquisition of upper jaw features upon Bosentan treatment reinforces the conserved role of the Dlx family in first arch morphogenesis. I focus on the comparison between chondrichthyans and mammals. The degree of similarities between chicken, mouse, and dogfish sharks in PA patterning may reflect their common ancestor, whereas the differences in expression can arise during evolution. In light of our findings, I proposed a model, shown below (Fig 3.20).

![Diagram of Dlx genes regulation](image)

**Fig. 3.20: A proposed model of Dlx genes regulation in md-PA**

Mesenchymal cells secrete the endothelin signalling molecule, Edn1, which activates Dlx5 and Dlx6 expression. It is proposed that Dlx3 and Dlx4 are downstream of Dlx5 and Dlx6. Thick orange arrows show the possibility of Dlx3 and Dlx4 regulation directly by the endothelin signalling pathway. Red lines show inhibition of Dlx1 and Dlx2 by endothelin.
Chapter 4
Assessing Candidate Dlx-Responsive Genes in PA1 of the Chick Embryo

Introduction

Dlx genes belong to the homeobox family and are evolutionary conserved. Dlx genes are homologues of Drosophila distal-less (Dll), and heavily impact the embryo’s development by playing a pivotal role in the forebrain, limb and craniofacial development (Anderson et al., 1997; Kraus & Lufkin, 2006; Kuwajima et al., 2006; Sajan et al., 2011; Stühmer et al., 2002). Dlx genes encode TFs that polarize PA1. Dlx5 and Dlx6 are particularly important for conferring mandibular identity to the “distal” half of PA1, and their knockout leads to the homeotic transformation of the tissue in mouse neonates and chicken (Acampora et al., 1999; Beverdam et al., 2002; Depew et al., 2005; Jeong et al., 2008; Kempf et al., 1998; Qiu et al., 1995, 1997; Robledo et al., 2002). Dlx5/6 null mice also show malformation of the palate and eyelids, revealing broader roles in craniofacial development. Though there are several Dlx5/6 transcriptional targets, confirmed by chromatin immunoprecipitation (ChIP), such as Large 1 (LARGE Xylosyl- And Glucuronyltransferase 1), Bmper and Msx1 in the inner ear (Sajan et al., 2011), little is known about Dlx5/6 targets in the lower jaw, although all six Dlx genes are expressed in PA1 in chick embryos (chapter 3).

Jeong et al. (2008) conducted a genome-wide profiling experiment in different combinations of Dlx5 and Dlx6 mutant mice in pharyngeal arches which enhanced our understanding of Dlx5- and Dlx6-responsive genes in mice. To compare wild-type and mutant Dlx5/6−/− embryos, Jeong and co-workers compared the transcriptional profiling of both groups at E10.5 mouse using microarray analysis. In total, they identified 63
differentially expressed genes, of which 39 were downregulated, and 24 were upregulated with a greater than 2-fold change. Twenty of these genes were not previously recognized as being potential targets of Dlx5 or Dlx6. I chose 5 of these 63 genes to test as Dlx targets in the chick embryo. In selecting these genes I placed an emphasis on those that have a jaw phenotype when knocked out in mice. For example, I did not select the Foxl2 gene, as the known craniofacial defect in loss-of-function mutants is in the eyelids, not in PA1 or in the derivatives of PA1 (Jeong et al., 2008). The five selected genes were Goosecoid (Gsc), Paired-like homeodomain transcription factor 1 (Pitx1), Heart and neural crest derivatives expressed transcript 1 (Hand1), Heart and neural crest derivatives expressed transcript 2 (Hand2) and Gastrulation brain homeobox 2 (Gbx2). All selected genes, with their known craniofacial defects in mice, are tabulated below (Table 4.1).

As the first criteria for target genes, I monitored the expression of each candidate gene during pharyngeal arch patterning to confirm overlap with Dlx5 and/or Dlx6 expression in md-PA1.

<table>
<thead>
<tr>
<th>Genes</th>
<th>Defects</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gsc</td>
<td>Malformed nose, ear, tongue, hypoplastic dentary and ectotympanic</td>
<td>(Rivera-pérez et al., 1995; Yamada et al., 1995)</td>
</tr>
<tr>
<td>Pitx1</td>
<td>Reduced mandibles, ectotympanic, clefted palate and bifurcated tongue missing gonial</td>
<td>(Lanctôt et al., 1999)</td>
</tr>
<tr>
<td>Hand1</td>
<td>Reduced mandibles, gonial and ectotympanic, cleft palate</td>
<td>(Barbosa et al., 2007)</td>
</tr>
<tr>
<td>Hand2</td>
<td>Distal truncation, midline and lower incisors fusion,</td>
<td>(H. Yanagisawa et al., 2003)</td>
</tr>
<tr>
<td>Gbx2</td>
<td>Reduced mandibles, otic capsule and middle ear ossicles</td>
<td>(Byrd &amp; Meyers, 2005)</td>
</tr>
</tbody>
</table>

Secondly, as a proxy for the specific loss of Dlx5 and Dlx6, we pharmacologically inhibited the endothelin signalling pathway at a critical window to down-regulate Dlx gene expression in md-PA1. Effects on the expression of the genes of interest were subsequently tested by WMISH and qPCR.
4.1 Results

4.1.1 Expression patterns in PA1

4.1.1.1 Gsc expression in chick pharyngeal arches

Gsc belongs to the bicoid subfamily of the paired homeobox family and encodes a TF that is a significant player in craniofacial and rib cage development in humans. The overall Gsc expression is summarized in Fig 4.1.

![Fig. 4.1: Gsc general expression summary](image)

(A) Schematic frontal view of a chick head depicting the expression of Gsc in PA1 (purple) and PA2 (yellow) (B) Gsc expression mapped onto the cap and hinge model of jaw patterning. In both cases the coloured areas represent an average of developmental stages between HH16-HH26.

At HH17, Gsc is expressed in PA1-PA4, highest in PA1, and progressively lower expression in more caudal arches, being only weakly detected in PA4. In each arch, expression is highest at the mesenchymal margins of the arch, being reduced or absent in the mesodermal core, and there is no expression in the surrounding ectoderm (Fig. 4.2 A-D). Viewed ventrally, expression in the mandibular arch extends in an arc that extends along the oral and aboral margin about halfway to the midline (Fig 4.2 C, D). Maxillary expression is restricted to the oral mesoderm. A similar pattern continues at HH18 (Fig. 4.2 E-H). Expression in the ectomesenchyme is obvious in PA2 where it surrounds the mesenchymal core (Fig. 4.2 C, D, G, H). At HH20 and HH22, the highest levels of Gsc expression were seen in the oral mesenchyme of both mx-PA1 and md-PA1, but not at
the midline (Fig. 4.2, K, L, O, P). Expression in PA2 retained the earlier pattern at HH20 (Fig. 4.2 K, L) but became restricted caudally at HH22 (Fig. 4.2 O, P), where it remained expressed until at least HH26 (Fig. 4.2 S, T). Between HH22 and HH26, the expression of Gsc in md-PA1 changed dramatically (Fig 4.2 Q-T). As md-PA1 elongated, expression can be seen in the distal-aboral region by HH26, in a mesenchymal domain that is almost complementary to the earlier expression. Faint expression was also detected in a thin medial zone in the ectomesenchyme that is oriented along the oral-aboral axis (Fig 4.2 T). Expression of Gsc in the otic vesicle also changes during this time, from being strongly expressed between HH17 and HH22 to being undetectable by HH26 (Fig 4.2 A, E, I, M, Q). Expression in the maxilla does not change so dramatically, continuing at the proximal-caudal margin of the mesenchyme at HH26 (Fig 4.2 S, T).
Fig. 4.2: Gsc expression in the chick embryo

In branchial arches, nasal placode, and otic vesicles by WMISH. The Hamburger-Hamilton stages are indicated. All images are shown from the right side. Each row represents different views of the same embryo.
4.1.1.2 *Pitx1* expression in chick pharyngeal arches

*Pitx1* also belongs to the bicoid subfamily and encodes a TF involved in limb and craniofacial development. The overall *Pitx1* expression is summarized in Fig 4.3.

![Diagram A](image1.png)  
**Fig. 4.3**: *Pitx1* general expression summary  
(A) Schematic frontal view of a chick head depicting the expression of *Pitx1* in PA1 (purple) (B) *Pitx1* expression mapped onto the cap and hinge model of jaw patterning. In both cases, the coloured areas represent an average of developmental stages between HH16-HH26.

The hindlimb-specific expression of *Pitx1* is apparent at all stages examined (Fig. 4.4 A, E, I, M, Q). At HH16, *Pitx1* is weakly expressed in the ectomesenchyme of PA1 - PA3, including the oral ectoderm of PA1, facial ectoderm and in an arc caudal to the eye (Fig 4.4 A-D). The same pattern continues at HH18, although expression in arches posterior to PA1 was undetectable (Fig 4.4 E-H). At HH20 *Pitx1* expression continues in the oral ectoderm of both PA1 branches, but the expression is now also apparent in the oral mesenchyme of md-PA1, restricted to a medial patch (Fig 4.4 I-L). This md-PA1 mesenchymal expression was reinforced by HH22 and the corresponding epithelial expression extends beyond the mesenchymal expression to the midline (Fig 4.4 M-P). At HH25 much of the distal half of md-PA1 is *Pitx1*-positive, with the exception of the most distal tissue at the midline (Fig 4.4 Q-T). Ectodermal expression in mx-PA1 was no longer detectable at HH25.
Fig. 4.4: *Pitx1* expression patterns of the chick embryo

*Pitx1* expression patterns of the chick embryo in the whole body, branchial arches, nasal placode, and otic vesicles by WMISH The Hamburger-Hamilton stage is indicated. All images are shown from the right side. Each row represents different views of the same embryo.
4.1.1.3 **Hand2 expression in chick pharyngeal arches**

Hand2, encoding a basic helix-loop-helix subfamily transcription factor, is involved in limb, pharyngeal arch and cardiac morphogenesis. The overall Hand2 expression is summarized in Fig 4.5.

![Fig. 4.5: Hand2 general expression summary](image)

(A) Schematic frontal view of a chick head depicting the expression of Hand2 in PA1 (purple) and PA2 (yellow) (B) Hand2 expression mapped onto the cap and hinge model of jaw patterning. In both cases, the coloured areas represent an average of developmental stages between HH16-HH26.

Expression in the posterior limb buds is a feature at all stages examined. At HH17 (Fig 4.6 A-D), Hand2 is expressed in PA1-PA4. In all arches, expression is restricted distally. In each arch, expression is highest at the mesenchymal margins of the arch, being reduced or absent in the mesodermal core, and there is no expression in the surrounding ectoderm. Expression levels are higher at HH19 (Fig 4.6 E-H) with no other change in the expression pattern. There is a gradual reduction in expression by HH21 (Fig 4.6 I-L) and at HH23 expression was no longer detected in PA3 or PA4 (Fig 4.6 M-P). At HH26, Hand2 expression is restricted to the most distal mesenchyme of md-PA1 but more broadly in the distal half of PA2 (Fig 4.6 Q-T).
Fig. 4.6: *Hand2* expression patterns of the chick embryo

*Hand2* expression patterns of the chick embryo in the whole body, branchial arches, nasal placode, and otic vesicles by WMISH. The Hamburger-Hamilton stage is indicated. All images are shown from the right side. Each row represents different views of the same embryo.
4.1.1.4 Hand1 expression in chick pharyngeal arches

Hand1 works in a complementary fashion with its parologue Hand2 and is also involved in cardiac and aortic arch morphogenesis. The overall Hand1 expression is summarized in Fig 4.7.

![Fig. 4.7: Hand1 general expression summary](image)

(A) Schematic frontal view of a chick head depicting the expression of Hand1 in PA1 (purple) and PA2 (yellow) (B) Hand1 expression mapped onto the cap and hinge model of jaw patterning. In both cases, the coloured areas represent an average of developmental stages between HH16-HH26.

Expression was largely undetectable in the pharyngeal arches of HH16 embryos (Fig 4.8 A-D) but was detected in the distal mesenchyme of PA1 and PA2 by HH18 (Fig 4.8 E-H). By HH20, Hand2 expression was expressed proximally in the ectoderm and distally in the mesenchyme of the first two arches (Fig 4.8 I-L), The ectodermal expression was largely absent by HH22 while expression in the distal arch mesenchyme continued, with expression being highest in the more caudal arches (Fig 4.8 M-P). Expression was reduced in PA1 and PA2 of HH23 embryos (Fig 4.8 Q-T).

4.1.1.5 Gbx2 expression in chick pharyngeal arches

The overall Gbx2 expression is summarized in Fig 4.9. At HH17 (Fig 4.10 A-D), Gbx2 expression can be detected proximally in PA1-PA4, highest in PA1 and progressively lower expression in more caudal arches, being only weakly detected in PA4. In each arch, expression is highest at the mesenchymal margins of the arch, being reduced or absent in the mesodermal core, there is no expression in the surrounding ectoderm. Viewed ventrally, expression in md-PA1 extends in a thin arc that extends along the oral margin
Fig. 4.8: Hand1 expression patterns of the chick embryo

Hand1 expression patterns of the chick embryo in the whole body, branchial arches, nasal placode, and otic vesicles by WMISH The Hamburger-Hamilton stage is indicated. All images are shown from the right side. Each row represents different views of the same embryo.
to the proximal aboral half. In PA2, the expression pattern of Gbx2 resembles Gsc, with transcripts detected in the ectomesenchyme that surrounds the mesenchymal core (Fig.4.10 A-D). Gbx2 is also expressed in the otic vesicle from HH17 to HH22.

![Diagram of Gbx2 expression](image)

**Fig. 4.9: Gbx2 general expression summary**

(A) Schematic frontal view of a chick head depicting the expression of Gbx2 in PA1 (purple) and PA2 (yellow) (B) Gbx2 expression mapped onto the cap and hinge model of jaw patterning. In both cases, the coloured areas represent an average of developmental stages between HH16-HH26.

By HH18, expression extends into the oral mesenchyme of mx-PA1 while expression in md-PA1 continues. In PA2, the expression pattern resembles that of Gsc, where it surrounds the mesenchymal core in a proximal domain (Fig.4.10 E-H). By HH20, Gbx2 expression continues in all four arches. In mx-PA1, expression in the oral mesenchyme continues and becomes progressively stronger to HH26. At HH20, expression was detected both orally and aborally in md-PA1 but this resolved to a largely oral domain that did not extend to the distal midline by HH22 (Fig. 4.10 I-P). Expression in PA2 mesenchyme resolves to a caudal-medial domain until at least HH26. The only md-PA1 expression changed significantly between HH22 and the last stage examined, at HH26. By HH26 Gbx2 expression remained in two discrete medially-positioned domains, one located orally and the other aborally (Fig. 4.10 S, T).
Fig. 4.10: Gbx2 expression profile

Gbx2 expression patterns of the chick embryo in the whole body, branchial arches, nasal placode, and otic vesicles by WMISH. The Hamburger-Hamilton stage is indicated. All images are shown from the right side. Each row represents different views of the same embryo.
4.1.2 Effect of endothelin disruption

Genes whose expression is dependent on Dlx5 or Dlx6 are predicted to be down-regulated when those genes are lost or their expression reduced. In the pharyngeal arches, such Dlx target genes should, in turn, be sensitive to disruptions in endothelin signalling. To disrupt the endothelin pathway, I took advantage of a known antagonist, Bosentan. I first monitored the effects of disrupting Edn signalling on the genes of interest by WMISH (Fig. 4.11).

We observed that Gsc expression in md-PA1 was downregulated but not entirely absent from either branch of PA1, whereas expression was lost entirely from PA2 and PA3 (Fig 4.11 A-D). Pitx1 expression was lost in the mesenchyme, but not in the oral ectoderm of md-PA1 in Bosentan-treated embryos. This highlights that Pitx1 expression in the oral ectoderm is not under the control of endothelin signalling (Fig 4.11 E-H). Hand2 expression was strongly down-regulated in PA1 and PA2, with some expression retained at the midline. PA3 and PA4 were less affected by Bosentan treatment (Fig 4.11 I-L). Although Hand1 was not expressed as strongly as Hand2 in md-PA1, expression was also downregulated in that tissue (Fig 4.11 M-P). In Bosentan-treated embryos, the main domains of Gbx2 expression in PA1 showed a variable response but were not consistently reduced compared to vehicle-treated controls (Fig 4.11 Q -T).

Next, I used quantitative RT-PCR to measure changes in gene expression. The mandibular branch of PA1 and the hyoid arch were dissected from three Bosentan- or vehicle-treated embryos and pooled (30 mg total wet tissue). Total RNA was extracted, and the transcripts were reverse-transcribed. PCR amplification of each gene was monitored with SYBR Green (Fig 4.12). Consistent with the in situ data, Gsc, Pitx1, and
Fig. 4.11: Differential expression of putative Dlx target gene

WMISH showing at 48 hr post-Bosentan treatment. Embryos were treated as previously described. Gsc control (A & C) and Bosentan-treated embryo (B & D); 3 of 4 embryos showed downregulation. Pitx1 control (E & G) and Bosentan-treated embryo (F & H); 4 of 4 embryos showed downregulation of Pitx1. Hand2 control (I & K) and Bosentan-treated embryo (J & L); 3 of 3 embryos showed downregulation. Hand1 control (M & O) and Bosentan-treated embryo (N & P); 4 of 4 embryos showed downregulation of Hand1. Gbx2 control (Q & S) and Bosentan-treated embryo (R & T); 2 of 4 embryos exhibited downregulation of Gbx2, whereas one embryo showed upregulation, and in one embryo, expression remained the same. LV, lateral view; FV, frontal view.
Hand2 was down-regulated in Bosentan-treated pharyngeal arch tissue, compared to vehicle-treated controls. The expression of Hand1 in Bosentan-treated arch tissue was highly variable. Unlike other genes I examined, there was not a consistent pattern of either up-regulation or down-regulation as both were seen in individual qPCR assays. While the graphed average suggests no change in Hand1 expression, the in situ data of individual embryos supports the notion that, like Hand2, Hand1 is also down-regulated when Edn signalling is compromised. Conversely, Gbx2 expression was up-regulated in treated arch tissue, by about 2.5-fold.

![Graph of gene expression levels](image)

**Fig. 4.12: RT-qPCR following Bosentan treatment**

Differential expression of candidate Dlx target genes upon inhibition of endothelin signalling pathway by Bosentan. Gene expression levels in md-PA1 and PA2 at HH21-22, 48 hr post-treatment. The whole mandibular process and hyoid arch were excised from treated and control embryos and pooled from 2-3 embryos. The gene of interest was normalized to SDHA expression. Control gene expression is set at 1 for each gene. Normalized expression is plotted relative to control tissue ±SEM. The experiment has been repeated a minimum of 6 times. The ΔCt between the GOI and the normalizer gene in each experiment was used to test for statistical significance using an unpaired Student’s t-test: *P<0.05, **P<0.01.
4.1.3 Effect of misexpressing Dlx5

Loss-of-function experiments have their limitations in identifying direct transcriptional targets. Furthermore, while Bosentan is a highly specific Ednra antagonist, I demonstrated that multiple Dlx genes directly or indirectly depend on endothelin signalling for the induction and/or maintenance of their expression during pharyngeal arch morphogenesis. Thus, I could not determine whether a Dlx protein directly regulated any of the genes whose expression was reduced, nor could we assign the effects to a specific parologue.

In contrast, gain-of-function experiments allow for testing individual paralogues for regulatory effects. To that end, I used an RCAS-Dlx5 virus to over-express Dlx5 in md-PA1 and ectopically express Dlx5 in mx-PA1. We injected Dlx5 virus particles at HH15-16 and collected them 24hr, 48hr and 72hr post-injection for WMISH for candidate target genes. The uninjected side of the embryo served as an internal control (Fig 4.13). While we reproducibly misexpressed Dlx5 in mx-PA1 using this strategy, we could not demonstrate ectopic expression of any candidate downstream target gene (Fig 4.13).

4.2 Discussion

We identified several genes whose expression in PA1 of mouse embryos is dependent on endothelin signalling and are therefore candidates for direct transcriptional regulation by Dlx5 and/or Dlx6. I monitored the expression of these genes in order to compare their expression patterns with Dlx family genes. These genes encode TFs and their dependence on endothelin signalling had previously not been documented in chick embryos. Together, this data positions Gsc, Pitx1, Hand2, and Hand1 downstream of the endothelin pathway and strengthens the case that Dlx proteins regulate them.
Fig. 4.13: Monitoring candidate gene expression following ectopic expression of *Dlx5*

(A) *Dlx5* expression in an *RCAS-Dlx5* infected embryo, visualized by WMISH. Note that virus injection was unilateral (yellow arrow) such that the contralateral (left) side of the embryo acts as an internal control. (B-E) WMISH for endogenous genes. (B) *Gsc*: 0 of 4 embryos showed ectopic expression. (C) *Pitx1*: 0 of 6 embryos showed ectopic expression. (D) *Hand2*: 0 of 4 embryos showed ectopic expression. (E) *Gbx2*: 0 of 5 embryos showed ectopic expression in *RCAS-Dlx5*-infected embryos. Yellow arrows point to expected sites of ectopic expression.

### 4.2.1 Comparison with orthologues

Comparing expression in other taxa facilitates insights into the development and evolution of vertebrate pharyngeal arches. We will compare the expression of *Gsc, Pitx1, Hand1, Hand2* and *Gbx2* with their orthologues in mouse pharyngeal arches, focusing on E10.5 mouse embryos and an equivalent chicken developmental stage (HH25-26).

*Gsc* expression in mouse and chicken follow the same expression pattern. In both organisms, *Gsc* is expressed in both branches of PA1. In mouse mx-PA1, expression is limited to the distal ectoderm and makes an arc from the oral to the aboral region (Jeong et al., 2008). Mouse *Gsc* expression in md-PA1 starts from the oral-distal region and extends to the aboral region through the midpoint of the proximal-aboral domain (Jeong et al., 2008).
The expression pattern of Pitx1 in chick and mouse md-PA1 is conserved, although the Pitx1 transcript in mx-PA1 may be relatively higher in mice compared to the chick (Jeong et al., 2008). In the chick, expression disappears from mx-PA1 at HH25, whereas, in mice, expression is still detected at an equivalent developmental stage (Jeong et al., 2008).

In mice, unlike chicks, Gbx2 is expressed only in the mandibular domain towards the proximal region, and in the distal-oral ectoderm, a weak expression is also observed in the ectomesenchyme (Jeong et al., 2008). Chicken Gbx2 is expressed in both branches of PA1 and, contrary to mouse Gbx2, chick Gbx2 transcripts are detected in the distal-proximal axis in md-PA1 at HH26.

Hand1 and Hand2 follow the same expression pattern not only between paralogues but between orthologues as well. In all arches, the expression of both genes is restricted distally.

4.2.2 Endothelin dependence of Dlx putative targets

Within PA1, not all gene expression domains are under endothelin or Dlx5/6 regulation. Furthermore, CNCCs colonization in PA1 and PA2 and its coordination with epithelium signalling to activate Dlx5 and Dlx6 is conserved in chicks and mice (Couly et al., 1993; Couly et al., 2002; Le Douarin et al., 1993; Le Douarin et al., 2004). I also noticed that blocking Edn signalling mostly affected the proximal territory of the mandibular prominence, where Dlx5 primarily resides in the mandibular arch at HH22-23. It can be hypothesized that in md-PA1, two sub-territories are differentially regulated by two or more signalling pathways; one is governed by endothelin, and the other might be via Fgf8 signalling (Vieux-Rochas et al., 2007, 2010), or by TFs other than the endothelin signalling effectors, Dlx5 and Dlx6. All of the genes I examined impact craniofacial development in their own right when knocked out in mice.

4.2.3 Gene expression overlaps with Dlx5/6 and other Dlx paralogue pairs

Next, I asked to what extent the expression domains of Gsc, Pitx1, Hand1, Hand2 and Gbx2 overlap with Dlx5 or Dlx6 or other Dlx gene expression in mdPA1. For most of the
developmental time points I examined, Dlx5 was expressed throughout most of the mandibular prominence. From HH16-HH26, Dlx6 expression largely overlaps with Dlx5 expression in the mandibular prominence. Gsc is expressed mainly in the proximal region, but at HH20 and HH22, it is expressed in distal-oral-ectoderm, and at HH26, expression shifts from oral to the distal-aboral ectoderm. Pitx primarily resides in the same territory as Dlx5 and Dlx6. Hand1 and Hand2 follow the same expression pattern and overlap in the same region as Dlx5 and Dlx6. All these findings are consistent with a model in which Dlx5 and Dlx6 contribute to regulating these genes. Notably, expression of Dlx3 and Dlx4 is also lost in Dlx5/6−/− mouse embryos and Bosentan-treated chicken embryos, so it cannot be ruled out that Dlx3 and/or Dlx4 contribute to the regulation of one or more of these genes. The endogenous expression domains and effects of disrupting endothelin signalling are schematized in Fig 4.14.
Fig. 4.14: Comparison of *Dlx* gene expression in mdPA1 with that of putative target genes

(A) Labelled frontal view of chick craniofacial primordia (B & C) Expression of *Dlx1* and *Dlx2* in magenta (D & E) in blue, expression of *Dlx3* and *Dlx4*. The relatively weaker expression of *Dlx3* at the midline is shown in a lighter colour. (F & G) expression of *Dlx5* and *Dlx6* in red. (H-L) Putative target gene expression in control and Bosentan-treated embryos at HH22-23. Gsc (H & H'), *Pitx1* (I & I'), *Hand2* (J & J') *Hand1* (K & K') *Gbx2* (L & L'). Orange (H- L) shows putative targets gene expression overlap with *Dlx1*+ *Dlx2* +*Dlx3*+ *Dlx5*+*Dlx6*. Increased *Gbx2* expression (L') is depicted with darker orange colour. Expression in mxPA1 is not shown for genes expressed there.
Gsc expression (Fig 4.14 H & H’) is reduced in Bosentan treated embryos. Gsc can potentially be a target of all Dlx proteins. But neither Dlx1 nor Dlx2 is downregulated when Edn signalling is inhibited, so Gsc cannot be the exclusive target of Dlx1 and Dlx2 in mdPA1. Furthermore, Gsc shares a common domain of expression with Dlx3 and Dlx4, so it can be hypothesized that Gsc might be a regulatory target of Dlx3 or Dlx4. Although the expression domain of Dlx4 is nested within the expression domain of Gsc which makes Gsc less likely to be a Dlx4 target. Pitx1 (Fig 4.14 I & I’) expression does not overlap with that of Dlx4, so Dlx4 cannot be the regulator of this paired like homeodomain gene. Like Pitx1, Hand1, Hand2, and Gbx2 (Fig 4.14 J-L’) share expression domains with Dlx1/2/3/5/6 but not with Dlx4. So they can be regulated by Dlx3, Dlx5 and/or Dlx6. Moreover, maxillary expression of Gbx2 remains unchanged when endothelin signalling is inhibited. Loss-of-function of any of these genes could not recapitulate the phenotype observed in Dlx5/6−/− embryos (Jeong et al., 2008), which likely puts Dlx5 and Dlx6 upstream in these regulatory pathways.

4.3 Conclusion

In this study, I profiled the expression of genes that were putative transcriptional targets of Dlx5 and/or Dlx6 in mouse md-PA1. I demonstrated that these genes are co-expressed to some extent with Dlx5 and Dlx6 between HH16 and HH25/26. The early activation of Dlx5 and Dlx6 may be responsible for regulating many genes that play a pivotal role in mandibular arch patterning, such as Gsc, Pitx1, Hand1, Hand2, and Gbx2, as evidenced by changes in the expression of these genes following antagonism of endothelin signalling and loss of both Dlx5 and Dlx6 expression. Therefore, these genes may be the targets of Dlx5 and/or Dlx6 in the chicken embryo. Our findings regarding endothelin signalling emphasize the evolutionary conservation of gene regulatory networks involved in patterning the PA1.
Chapter 5

A computational Approach to cis-Regulatory Element Discovery in the Chick: Identifying Enhancers for Cranial Neural Crest and/or the First Pharyngeal Arch

Introduction

The journey from fertilized cell to complete organism encompasses many complex developmental pathways, morphogenesis and organogenesis. Spatiotemporal gene expression is at the core of this process. Transcription of a gene starts when a TF binds to cis-regulatory elements (CRE) via individual binding sites, also known as response elements (RE). Physical interaction between the RE and the TF depends on the DNA binding specificity of the TF and interactions with other CRE-bound proteins (Serfling et al., 1985). Enhancers are one kind of CRE and may be present from hundreds of base pairs (bp) to hundreds of kilobases (kb) or more from the regulated gene. They are varied in size, from 10 bp to 1000 bp in length. (Blackwood & Kadonaga, 1998; Yáñez-Cuna et al., 2013) and can be upstream or downstream of a transcription start site or part of a transcription unit (Heinz et al., 2015).

Enhancers play a pivotal role in the target gene’s transcription by physical interaction with the proximal promoter. The physical interaction can be confirmed using a variety of methods, including formaldehyde crosslinking of DNA and TFs, chromosome conformation capture (3C), and fluorescence in situ hybridization (FISH) to demonstrate the spatial proximity of enhancers and promoters in the nucleus (Deng et al., 2012; Deng et al., 2014). The enhancer-promoter interaction usually facilitates the delivery or stability of the pre-initiation complex (PIC), RNA polymerase II (pol II), and promoter-bound TFs in a time and tissue-specific manner (Deng et al., 2012).
An enhancer’s state can be classified as active, inactive, primed, poised, or bivalent (Ernst & Kellis, 2010). These different states can be recognized by the combination of methylation and acetylation of histone proteins. Inactive enhancers are buried in a compact nucleosome region. Active enhancers have epigenetic marks like histone 3 lysine 27 acetylation (H3K27ac), and the TF(s) that are necessary for regulated gene expression are bound. Primed enhancers are like active enhancers, but they do not produce enhancer RNA (eRNA). They have the same epigenetic marks as active enhancers but need an additional cue in the form of signal transduced peptide, adaptor or coactivator signal (eRNA) to become active. Primed enhancers have epigenetic marks such as H3K4me1 and H3K4me2 (Eliezer & Joanna, 2013). Poised enhancers are like primed enhancers, but they are distinguished by the presence of the H3K27me3 mark (Eliezer & Joanna, 2013). These modifications need to be removed to allow them to become active enhancers. A “bivalent” modification represents an overlapping presence of both permissive H3K4me3 and repressive H3K27me3 histone modifications at the same locus. Interestingly, all Dlx gene family members and genes discussed in chapter 4 (Gsc, Pitx1, Gbx2 and Hand2) have bivalent domains in human fetal craniofacial tissue; Hand1 is an exception (Wilderman et al., 2018). In general, enhancers display enrichment of H3K4me1 or H3K4me2 and depletion of H3K4me3 compared with promoters (Heintzman et al., 2007).

It is believed that metazoan chromosomes are composed of different topologically-associated domains (TAD) (Dixon et al., 2016; Dekker & Heard, 2015). These TADs have dozens of transcriptional units and multiple enhancers (Furlong & Levine, 2018). In most cells, CTCF and cohesion protein mediates enhancer selection among the enhancers present in the TAD (Wendt et al., 2008; Splinter et al., 2006). In the case of multiple enhancers in a TAD domain, the enhancer has specificity for a binding motif sequence in the promoter, the TATA box or downstream promoter element (DPE) (Zabidi & Stark, 2019). Interestingly, some enhancers can interact with all promoters close to them (Symmons et al., 2014). In mouse models, for example, major facilitator superfamily-related transporter 4 (Mrf4) and major facilitator superfamily-related transporter 5 (Mrf5) have a common enhancer, but their expression pattern is different due to the
heterologous promoter. These promoter sequences display differential specific residues at the enhancer-promoter interphase to express one gene or other linked genes (Chang et al., 2004). Generally, it is believed that one proximal promoter has a more significant competitive advantage over the distal one and does this by sequestering the enhancer from the other promoter (Furlong & Levine, 2018; Haberle & Stark, 2018).

Enhancers and promoters may interact through tracking, linking and looping. Evidence of the proposed tracking mechanism came from inserting an insulator between the enhancer and promoter (Kong et al., 1997). In tracking, pol II and the TF assemble at the enhancer region. Pol II processes along with the DNA and pulls the enhancer with it. This mechanism is possible over short distances (Furlong & Levine, 2018). Another short-distance mechanism of enhancer-promoter interaction is linking. In linking, enhancer-bound TFs increase their concentration through dimerization or oligomerization and transcribe the target gene. The *Drosophila* ChIP protein may enhance transcription in this way (Morcillo et al., 1997). Studies on the LIM-domain binding 1 (Lbd1) protein suggest that it can mediate loop formation (Deng et al., 2012). Regulatory TFs and basal transcriptional machinery assemble on the enhancer and promoter in the looping model, respectively. Interaction occurs via the looping of intervening DNA. This bending is powered by adenosine triphosphate. Evidence of looping first came from *E. coli* over 30 years ago (Dunn et al., 1984). It is also believed that when the enhancer is very far away (100 kb – 1 Mb), tracking, linking and looping may all play a role in facilitating the transcription of a target gene.

There appear to be at least two kinds of enhancers when it comes to the interaction of TFs with enhancer sequences. The term enhanceosomes describe enhancers in which the relative order and orientation of the TF binding site are essential, and all TF binding and outputs are integrated, and they act as one unit (Thanos & Maniatis, 1995). Given the stringent requirement for sequence identity over the entire enhancer and the sensitivity to mutations, orthologous enhancers of this kind can be identified by a remarkable level of sequence conservation over long evolutionary distances.
On the other hand, the billboard model describes a group of enhancer sequences that are more flexible with respect to the number, order, and orientation of the TF binding sites. In this latter kind of enhancer, TF binding sites can be scrambled experimentally or naturally over evolutionary timescales without significant effects on enhancer function. Enhancers of this kind will be difficult, if not impossible, to identify by sequence conservation alone since the function of the enhancer can be maintained even if the larger sequence context of the enhancer is not conserved (Arnosti & Kulkarni, 2005).

One of the more remarkable features of developmentally active enhancer elements is their strong evolutionary sequence conservation, which can be observed in species both distantly and closely related (Dickel et al., 2013). Guided by the concept that enhancer elements exist within homologous regions conserved between the genome of various species, the identification of complex developmentally active enhancer elements has shifted from a blind approach towards a more effective approach using comparative genomics (Visel et al., 2009). In addition to comparative genomics, ChIP-sequencing of a particular molecular, protein or histone marker is frequently bound within enhancer regions, such as cofactor p300. P300 and its paralogue CREB binding protein (CBP) play an essential role in gene regulation by acetylating histones in DNaseI hypersensitive regions (C. Chrivla, Roland S. Kwok. L, Masatohl.H, 1993; Wang et al., 2009). Their genomic occupancy is the signature of transcriptional enhancers (Heintzman et al., 2007) and has been used as a tool to identify transcriptional regulatory sequences (Ong & Corces, 2011). Through these approaches, and with the availability of computational tools and resources, researchers can now more accurately identify and map candidate enhancer regions involved during development and use this information to then direct future experimentation in vivo.

This study used computational methods to identify conserved non-coding sequences in the chicken genome that were likely to have enhancer activity in PA1. For this purpose, computational resources and p300 ChIP-seq data were combined in an attempt to identify conserved enhancer regions in the chick genome with putative function during the development of the beak. I then sought to validate these sequences following in ovo
electroporation of reporter plasmids. Next, I asked if Dlx5 and or Dlx6 regulate these sequences.

5.1 Results

5.1.1 *In silico* analysis of overlapping VISTA Enhancer and p300 ChIP-Seq data

To address the need for an enhanced understanding of transcription regulation in chicken embryogenesis, we sought to identify chick enhancers for genes expressed in PA1. Our approach was first to identify a high-quality set of mouse sequences using public genomic resources and then locate as many as possible in the chick genome via sequence identity. We first defined the intersection of two mouse data sets: PA1-active enhancers (VISTA) found in the Face–Base Hub. FaceBase compiles data from ChIP and Illumina platform sequencing of enhancer-associated p300 proteins within the dissected craniofacial tissue of mouse embryos at E11.5 (Visel et al., 2009). Putative enhancer regions validated using transgenic mouse assays with a reproducible pattern at E11.5 were assembled via VISTA Enhancer Browser and mapped within the Mouse GRCm38/ genome (Visel et al., 2009). The resulting database containing the p300 DNA-associated peak was exported from FaceBase and mapped to the Mouse GRCm38/mm10 genome (Fig. 5.1).

Criteria for reproducible enhancer expression were according to VISTA Enhancer Browser but also reflected the limited sample size found in resulting *in vivo* experimentation (Visel et al., 2009). The Face-Base Hub utilizes the p300 binding sites in the mouse genome at E11.5. 470 enhancer regions in the mouse genome aligned with the p300 ChIP-seq data (see Fig. 5.2). 367 out of the 470 (about 78%) enhancers were not active in pharyngeal arch tissue. The remaining 103 enhancers (22%) showed activity in PA tissue (Fig. 5.2). Only 85 out of 103 (82 %) of these showed a reproducible pattern in PA, whereas the others failed to show reproducibility (Fig. 5.2). Thirty-two of the 85 mouse enhancers were within 500 kb of a gene that has been expressed in PA1 in mouse embryos, but only 7 of the 32 enhancers satisfied the criteria for sequence conservation in the chicken genome i.e; >80% identity over at least 100 bp.
Fig. 5.1: A general scheme of the prediction method

Regions were identified from overlapping FaceBase p300 ChIP-seq peak sequences (upper left) and VISTA Enhancer validated enhancer sequences (lower left) mapped along the Mouse GRCm38/mm10 genome of the UCSC Genome Browser (right).

One representative overlap of the p300 ChIP seq. enhancer from Face-Base hub and VISTA Enhancer, validated in the transgenic mouse, is shown in Fig 5.3. Lastly, enhancer regions cloned from the mouse genome with reproducible PA-specific activity were further refined based on their proximity to genes that were expressed within PA1 of mouse embryos. Genes located within 500 kb of each candidate enhancer region were examined for expression in PA1 using mouse and chick gene expression atlas.

All 7 regions are shown in Table 5.1 with their coordinates in the mouse and their equivalent DNA-aligned fragments in the chicken genome. Five of them were cloned while repeated attempts to amplify mm426 and mm466 from gDNA were unsuccessful.
5.1.2 In ovo putative enhancer testing

The predicted enhancer regions were cloned upstream of the thymidine kinase promoter in ptk-EGFP version 2 (Kamachi et al., 2009; Uchikawa, 2008). After cloning, the orthologous chicken genomic sequences were given gg prefixes attached to the VISTA designation, hence gg384, gg901, gg1090, gg1114 corresponding to mm384, mm901, mm1090, and mm1114 VISTA enhancer elements respectively. Electroporation of reporter plasmids at the midbrain-hindbrain at HH9-10 allows for the monitoring of enhancer activity in migrating NCCs or PA1 (Betancur et al., 2010a Betancur et al., 2011). We first performed a series of efficiency-viability tests to test find conditions for both acceptable transfection efficiency and embryo viability (Fig. 5.4). HH9-10 chicken embryos were microinjected with a pCAGGS-EGFP plasmid solution and electroporated at 15V for a variable number of pulses. While embryo viability did not vary significantly with the number of pulses, administration of 2 pulses resulted in low-level GFP expression in 63% of surviving embryos.
Fig. 5.3: A representative overlap mapping of p300 enhancer and VISTA enhancer in UCSC genome browser

Sample sequence conservation output for mm924, a putative Runx2 enhancer region. Sequence conservation determined by (A) BLAT alignment in Chicken GRCg6a/galGal6 genome; (B) mVISTA glocal alignment display of Mouse GRCm38/mm10 putative mm924 enhancer sequence (x-axis) and homologous Chicken GRCg6a/galGal6 sequence (y-axis), with pink peaks corresponding to >100bp nucleotide alignment of over 70% identity within the gene’s conserved non-coding sequence (CNS); and (C) UCSC Genome Browser alignment display of mm924 (top), wiggle track display of nucleotide conservation scores between the 60 aligned vertebrate genomes (phastCons and phyloP) (middle) and display of multiple pairwise alignments (multiz) amongst model vertebrate species, including the chicken (bottom).
Table 5.1: All VISTA and Face-Base enhancers overlapping in craniofacial mouse tissue

All overlapped enhancers with their element number in VISTA Browser and rank number in Face-Base are shown with their coordinates, closest gene, their distance from the gene which is involved in craniofacial development, and coordinates in both the mouse as well as aligned chicken coordinates in the latest (GRCg6a/gal Gal6) chicken assembly. Percentage similarity within the DNA fragment between the mouse and chicken is also computed.

<table>
<thead>
<tr>
<th>Enhancer Name (VISTA/FaceBase)</th>
<th>Gene</th>
<th>Mouse GRCm38/mm10 Coordinates</th>
<th>Distance from Gene</th>
<th>Chicken GRCg6a/galGal6 Coordinates</th>
<th>Distance from Gene</th>
<th>Enhancer Alignment Between Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>mm1090/Rank-479</td>
<td>Prdm16</td>
<td>chr4:154,801,902-154,803,150</td>
<td>165 kb</td>
<td>chr21:1351349-1352131</td>
<td>within prdm16</td>
<td>89.6%</td>
</tr>
<tr>
<td>mm426/Rank-2104</td>
<td>Msx1</td>
<td>chr5:37825669-37829220</td>
<td>within intron</td>
<td>chr4:78802377-78805012</td>
<td>within intron</td>
<td>82.0%</td>
</tr>
<tr>
<td>mm1114/Rank-4759</td>
<td>Spry2</td>
<td>chr14:105979373-105980837</td>
<td>83 kb</td>
<td>chr1:154020273-154020884</td>
<td>34.5 kb</td>
<td>89.2%</td>
</tr>
<tr>
<td>mm924/Rank-1371</td>
<td>Runx2</td>
<td>chr17:44617959-44619377</td>
<td>within intron</td>
<td>chr3:109434210-109434533</td>
<td>within intron</td>
<td>89.8%</td>
</tr>
<tr>
<td>mm901/Rank-2101</td>
<td>Twist2*</td>
<td>chr1:91962929-91965960</td>
<td>114.9 kb</td>
<td>chr7:6244147-6245429</td>
<td>56.8 kb</td>
<td>89.5%</td>
</tr>
<tr>
<td>mm384/Rank-2424</td>
<td>Bmp4*</td>
<td>chr14:45943841-45945127</td>
<td>438.4 kb</td>
<td>chr5:58582275-58582692</td>
<td>131.9 kb</td>
<td>91.8%</td>
</tr>
<tr>
<td>mm466/Rank-1829</td>
<td>Snai2*</td>
<td>chr16:15084593-15085708</td>
<td>375.3 kb</td>
<td>chr2:108256017-108256719</td>
<td>172.2 kb</td>
<td>87.4%</td>
</tr>
</tbody>
</table>

*Not the closest transcription start site with respect to the putative enhancer region.
Fig. 5.4: Electroporation of pCAGGS-EGFP control plasmid in HH9-10 chicken embryos

A number of biological replicates underwent embryo were microinjected and electroporated at 15V for 2, 3, or 5 pulses. The survival rate of embryos treated with 2, 3, and 5 pulses was calculated 24 hours and 48 post electroporation. (B) GFP expression in surviving embryos after 48h.

In contrast, all surviving embryos showed GFP-positive NCC after 3 pulses so this became the standard protocol. Extensive delivery of pCAGGS-mCherry to NCC that subsequently formed migratory streams into PA1-4 further validated the stage and site of electroporation (Fig. 5.5). Using pCAGGS-EGFP and pCAGGS-mCherry plasmids, I demonstrated that co-injected plasmids were co-expressed following electroporation and that we could reproducibly target CNCC at HH9-10 (Fig 5.6).
Fig. 5.5: Chicken embryo demonstrating NCC migration into pharyngeal arches

pCAAGS-mCherry was electroporated at the junction of the midbrain and hindbrain at HH9. Red arrows point to migratory streams of mCherry-expressing NCC at HH13 (A) or HH16 (B).
Fig. 5.6: Co-expression of reporter plasmids in cranial neural crest

3mg/ml of each pCAGGS-GFP and pCAGGS-mCherry plasmids were co-injected into the lumen of the hindbrain of an HH10 chick embryo and electroporated on the left side. (A) Bright-field image of the ex-vivo embryo, 24 hrs post-electroporation. Yellow arrows point to the maxillary and mandibular branches of the first pharyngeal arch (PA) and neural tube (NT). (B) EGFP expression. (C) mCherry expression. (D) The merged image shows overlapping fluorescence in scattered migrating NCC and the dorsal midbrain and more caudal neural tube. Yellow arrows in B show the NCCs residing in PA1.
Next, I screened all genomic sequences in an orientation we deemed to be native, such that the orientation of the genomic sequence with respect to EGFP transcription was the same as its orientation with respect to transcription of the nominal regulated gene shown in Table 5.1. Each sequence was tested in this “forward” orientation at 24 hr and 48 hr post-electroporation. I also tested each sequence in the opposite orientation at 24 hr post-electroporation.

Fig 5.7A shows the genomic orientation of the gg901 putative enhancer element (mm901 VISTA enhancer element). This enhancer region is in intron 16 of the HDAC4 gene and 56.8 kb downstream of the Twist2 gene in the chicken genome, which plays a role in osteoblast maturation and development during craniofacial patterning. Fig 5.7 B & C depicts the native and “antisense” orientation of the enhancer element in a ptk-EGFP plasmid, respectively.

**Fig. 5.7: Genomic landscaping and cloning strategy of gg901**

**(A)** The genomic landscape of element gg901. The orange triangle represents the location of the elements in the context of the chicken genome. Two vertical blue lines show the distance between the middle of the cloned sequence and the TSS of the Twist2 gene. **(B)** Element cloned in a “sense” or native orientation (+) between XhoI and KpnI restriction site. A triangle with an apex up is used to show that the element is downstream of the presumptive endogenous regulated gene. **(C)** Element cloned into the reporter plasmid in the reverse (-) orientation, depicted by an apex-down triangle.
The gg901 reporter was active at 24 hr (Fig 5.8) with reporter activity in PA1, PA3, and PA4 in one of two surviving embryos. At 48 hr post-electroporation, one of the two surviving embryos showed very few scattered GFP-positive cells (Fig 5.9). Similarly, when tested in the opposite orientation (Fig 5.10), one embryo out of three embryos was positive for GFP reporter activity at 24hr.
**Fig. 5.8: In ovo activity of gg901 at 24 hr post-electroporation**

(A-D) Embryos were injected and electroporated with gg901-tk-EGFP plasmid (A) gg901 genomic sequence activity in which GFP cells are visible in migrating NCC. Yellow arrows point to the GFP-positive cells (B) pCAAGS-driven RFP, reporting the extent of electroporation (C) Bright-field image (D) Merged image.
Fig. 5.9: In ovo activity of gg901 at 48 hr post-electroporation

(A-D) Embryos were injected and electroporated with gg901-tk-EGFP plasmid (A) Bright-field image (B) gg901 genomic sequence activity in which GFP cells are visible in migrating NCC. Yellow arrows point to the GFP-positive cells (C) pCAAGS-driven RFP, reporting the extent of electroporation (D) Merged image.
Fig. 5.10: In ovo activity of “antisense” gg901 at 24 hr post-electroporation

(A-D) Embryos were injected and electroporated with gg901rev-tk-EGFP plasmid (A) Bright-field image (B) gg901 genomic sequence activity in which GFP cells are visible in migrating NCC. The yellow arrow points to the GFP-positive cells (C) pCAAGS-driven RFP, reporting the extent of electroporation (D) Merged image.
Fig 5.11 A shows the genomic orientation of gg384. This enhancer region is flanked between \textit{Loc1070} and \textit{DDHD1} genes and 132 kb downstream of \textit{Bmp4}. \textit{Bmp4} is the closest gene involved \textit{in vivo} in cartilage formation. Fig 5.11 B & C show the native and opposite orientation of the enhancer element in a ptk-EGFP plasmid, respectively.

![Diagram](image)

**Fig. 5.11: Genomic landscaping and cloning strategy of gg384 element**

(A) The genomic landscape of element gg384. The orange triangle represents the location of the elements in the context of the chicken genome. Two vertical blue lines show the distance between the middle of the cloned sequence and the TSS of the \textit{Bmp4} gene. (B) Element cloned in a “sense” or native orientation (+) between XhoI and KpnI restriction site. A triangle with an apex up is used to show that the element is downstream of the presumptive endogenous regulated gene. (C) Element cloned into the reporter plasmid in the reverse (-) orientation, depicted by an apex-down triangle.

Fig 5.12 shows the activity of gg384 at 24hr post-injection-electroporation. At this time point, only three embryos were alive. The gg384-reporter was correspondingly active in one embryo at 24hr (Fig 5.12 A-D) with a reporter activity in a small number of cells as compared to the overall electroporation efficiency.
Fig. 5.12: In ovo activity of gg384 at 24 hr post-electroporation

(A-D) Embryos were injected and electroporated with gg384-tk-EGFP plasmid. (A) Bright-field image (B) gg384 genomic sequence activity in which GFP cells are visible in migrating NCC. Yellow arrows point to the GFP-positive cells (C) pCAAGS-driven RFP, reporting the extent of electroporation (D) Merged image.
At 48 hr post-electroporation, two embryos survived but neither had any GFP-positive cells, despite correctly targeted electroporation (Fig 5.13). Similarly, when tested in the opposite orientation (Fig 5.13 E-H), two of the two embryos were negative for GFP reporter activity at 24 hr.

The gg384 genomic construct was electroporated into the neural tube region, which targets the migrating neural crest cells at a 3 mg/ml concentration along with RFP. RFP is expressed from the control plasmid, and its expression is driven pCAGGS-mCherry reporter plasmid which was under the control of the chick β-actin promoter, indicating the location of successful electroporation. A-D denotes candidate enhancer element gg384 images at 48hr post-injection and electroporation in the sense direction; lastly, E-H denotes candidate enhancer element gg384 images at 24hr post-injection and electroporation in the antisense or opposite direction.

Fig 5.14 A shows the genomic orientation of gg1114. This genomic region is flanked by Loc1125 and Spry2 and is 34.4 kb upstream of Spry2, which is involved in eye lens development. Fig 5.14 B & C show the native and opposite orientation of the putative enhancer element in a ptk-EGFP plasmid, respectively.
Fig. 5.14: Genomic landscaping and cloning strategy of gg1114 element

(A) The genomic landscape of element gg1114. The orange triangle represents the location of the elements in the context of the chicken genome. Two vertical blue lines show the distance between the middle of the cloned sequence and the TSS of the spry2 gene. (B) Element cloned in a “sense” or native orientation (+) between XhoI and KpnI restriction site. A triangle with an apex up is used to show that the element is downstream of the presumptive endogenous regulated gene. (C) Element cloned into the reporter plasmid in the reverse (-) orientation, depicted by an apex-down triangle.

The gg1114-reporter was active at 24 hr (Fig 5.15 A-D) with reporter activity in one surviving embryo out of three. At 48 hr post-electroporation, only two embryos survived (Fig 5.15 E-H). One embryo showed two GFP-positive cells (Fig 5.15 F), but this was considered an artifact since many cells had been electroporated, and the “cells” were not in a region of interest. After repeating the experiment, an additional two embryos showed no GFP reporter activity. Similarly, when tested in the opposite orientation (Fig 5.15 I-L), all three embryos were negative for GFP reporter activity at 24 hr.
Fig. 5.15: In ovo candidate enhancer gg1114 screening

EGFP reporter detection in ovo. Embryos were injected and electroporated with EGFP plasmids containing gg384 putative enhancers and pCAGGS-ERFP reporter plasmid which was under the control of chick β-actin promoter. At HH10 electroporation was done into the hindbrain region, which targets the migrating neural crest cells. A-D refers to candidate enhancer element gg1114 images at 24hr post-injection and electroporation in the sense or native direction; E-H denotes candidate enhancer element gg1114 images at 48hr post-injection and electroporation in the sense direction; lastly, I-L denotes candidate enhancer element gg1114 images at 24hr post-injection and electroporation in the antisense or opposite direction.

Fig 5.16 A shows the genomic orientation of the gg924. This sequence is between the 5th and 6th exon (5th intron) of Runx2, which plays a vital role in osteoblast differentiation and skeletal morphogenesis. Fig 5.16 B & C shows the native and opposite orientation of the enhancer element in a ptk-EGFP plasmid, respectively.
Fig. 5.16: Genomic landscaping and cloning strategy of gg924 element

(A) The genomic landscape of element gg924. The orange triangle represents the location of the elements in the context of the chicken genome. This genomic sequence is located in the introns of the Runx2 gene. (B) Element cloned in a "sense" or native orientation (+) between XhoI and KpnI restriction site. A triangle with an apex up is used to show that the element is downstream of the presumptive endogenous regulated gene. (C) Element cloned into the reporter plasmid in the reverse (-) orientation, depicted by an apex-down triangle.

Fig 5.17 shows a representative embryo, 24 hr post-electroporation with pgg924-tk-EGFP. No reporter activity was seen in the three surviving embryos. At 48hr post-electroporation, only two embryos survived (Fig 5.17 E-H) with no reporter activity in either. Similarly, when tested in the opposite orientation (Fig 5.17 I-L), both surviving embryos were negative for GFP reporter activity at 24 hr post-electroporation.

Fig 5.18 A shows the genomic orientation of the gg1090 element. This putative enhancer is between the 1st and 2nd exon of the prdm16 gene. The exact craniofacial function of prdm16 is unknown, but it is a transcriptional regulator in many cells and is expressed in the forebrain and pharyngeal arches in the chick. Fig 5.18 B & C shows the native and opposite orientation of the putative enhancer element in a ptk-EGFP plasmid, respectively.
The gg924 genomic construct was electroporated into the neural tube region, which targets the migrating neural crest cells at a 3 mg/ml concentration along with RFP. RFP is expressed from the control plasmid, and its expression is driven pCAGGS-ERFP reporter plasmid which was under the control of chick β-actin promoter, indicating the location of migrating NCC and successful electroporation. A-D refers to candidate enhancer element gg924 images at 24hr post-injection and electroporation in the sense direction; E-H denotes candidate enhancer element gg924 images at 48hr post-injection and electroporation in the sense direction; lastly, I-L denotes candidate enhancer element gg924 images at 24hr post-injection and electroporation in the antisense direction.

**Fig. 5.17: In ovo candidate enhancer gg924 screening**
**Fig. 5.18: Genomic landscaping and cloning strategy of gg1090 element**

(A) The genomic landscape of element gg1090. The orange triangle represents the location of the elements in the context of the chicken genome. The genomic sequence is in the introns of the \textit{prdm16} gene. (B) Element cloned in a “sense” or native orientation (+) between Xhol and KpnI restriction site. A triangle with an apex up is used to show that the element is downstream of the presumptive endogenous regulated gene. (C) Element cloned into the reporter plasmid in the reverse (-) orientation, depicted by an apex-down triangle.

Fig 5.19 A-D shows a representative embryo, 24 hr post-electroporation with pgg1090-tk-EGFP. No reporter activity was seen in seven surviving embryos. At 48 hr post-electroporation, only one embryo survived (Fig 5.19 E-H) with no GFP activity. Similarly, when tested in the opposite orientation (Fig 5.19 I-L), the only surviving embryo was negative for GFP reporter activity at 24 hr.

Fig 5.20 A shows the genomic orientation of the \textit{ggHand2} putative enhancer, corresponding to the \textit{Hand2} mouse enhancer evaluated by Charité \textit{et al.} (2001). This enhancer region is present between \textit{Loc10705} and \textit{Loc10174} and 4.7 kb \textit{upstream} of the \textit{Hand2} gene in the chick genome, \textit{Hand2} is involved in pharyngeal arch morphogenesis. Fig 5.20 B & C shows the native and opposite orientation of the enhancer element in a ptk-EGFP plasmid, respectively.
Fig. 5.19: In ovo candidate enhancer gg1090 screening

The gg1090 genomic construct was electroporated into the neural tube region, which targets the migrating NCCs at a 3 mg/ml concentration and RFP. A-D refers to candidate enhancer element gg1090 images at 24hr post-injection and electroporation in the native direction; E-H denotes candidate enhancer element gg1090 images at 48hr post-injection and electroporation in the native or sense direction; lastly, I-L denotes candidate enhancer element gg1090 images at 24hr post-injection and electroporation in the opposite direction.
Fig. 5.20: Genomic landscaping and cloning strategy of ggHand2 element

(A) The genomic landscape of element ggHand2. The orange triangle represents the location of the elements in the context of the chicken genome. Two vertical blue lines show the distance between the middle of the cloned sequence and the TSS of the Hand2 gene. (B) Element cloned in a “sense” or native orientation (+) between XhoI and KpnI restriction site. A triangle with an apex up is used to show that the element is downstream of the presumptive endogenous regulated gene. (C) Element cloned into the reporter plasmid in the reverse (-) orientation, depicted by an apex-down triangle.

Fig 5.21 A-D shows a representative embryo, 24 hr post-electroporation with pggHand2-tk-EGFP. No reporter activity was seen in the two surviving embryos. At 48 hr post-electroporation (Fig 5.21 E-H) and in the opposite orientation, 24hr post-electroporation (Fig 5.21 I-L), two embryos survived of each with no obvious enhancer activity in any.
The ggHand2 enhancer construct was electroporated into the neural tube region, which targets the migrating neural crest cells at a 3 mg/ml concentration along with RFP. RFP is expressed from the control plasmid, and its expression is driven pCAGGS-ERFP reporter plasmid which was under the control of chick β-actin promoter, indicating successful electroporation. A-D refers to candidate enhancer element ggHand2 images at 24hr post-injection and electroporation in the sense direction; E-H denotes candidate enhancer element ggHand2 images at 48hr post-injection and electroporation in a native; lastly, I-L denotes candidate enhancer element ggHand2 images at 24hr post-injection and electroporation in the antisense direction.

Fig. 5.21: In ovo candidate enhancer ggHand2 screening
Table 5.2: Number of GFP-positive embryos at different time windows

GFP-positive count and orientation in PA1 were tabulated also compared with No. of mouse embryos in VISTA enhancer browser, which showed activity of the same enhancer in PA1.

<table>
<thead>
<tr>
<th>Enhancer elements ID</th>
<th>Orientation</th>
<th>No. positive at 24hpe</th>
<th>No. positive at 48hpe</th>
<th>No. positive at 24hpe</th>
<th>No. positive in VISTA Browser</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Forward</td>
<td>Forward</td>
<td>Reverse</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>gg901</td>
<td>1/2</td>
<td>1/2</td>
<td>1/3</td>
<td></td>
<td>5/5</td>
</tr>
<tr>
<td>gg384*</td>
<td>1/3</td>
<td>0/2</td>
<td>0/2</td>
<td></td>
<td>8/9</td>
</tr>
<tr>
<td>gg1114</td>
<td>1/3</td>
<td>1/2</td>
<td>0/3</td>
<td></td>
<td>3/4</td>
</tr>
<tr>
<td>gg924</td>
<td>0/3</td>
<td>0/2</td>
<td>0/2</td>
<td></td>
<td>6/8</td>
</tr>
<tr>
<td>gg1090</td>
<td>0/7</td>
<td>0/1</td>
<td>0/1</td>
<td></td>
<td>8/8</td>
</tr>
<tr>
<td>ggHand2</td>
<td>0/2</td>
<td>0/2</td>
<td>0/2</td>
<td></td>
<td>N/A</td>
</tr>
</tbody>
</table>

*Very small number of GFP-positive cells compared to the electroporated tissue

5.1.3 Identification of Dlx regulated transcriptional enhancer upstream of Gsc

Gsc was strongly down-regulated in the absence of Dlx5 and Dlx6 expression following Bosentan treatment (Fig. 4.11 B & D). To identify conserved non-coding sequences in the Gsc locus that might represent Dlx-responsive regulatory elements, I performed an in silico analysis of the Gsc locus. Using the Evolutionary Conserved Regions (ECR) Browser, I identified a 0.6kb highly conserved region (between mammals and birds) with putative Dlx5/6 binding sites, (A/C/G)TAATT(G/A)(C/G), as previously determined in a binding assay (Feledy et al., 1999). This sequence is located 3.4 kb upstream of Gsc (Fig 5.22 A & B).
**Fig. 5.22:** Identification of a Dlx-regulated enhancer upstream of Gsc in chick genome

(A) Evolutionary conservation of the genomic sequence upstream of Gsc analyzed using ECR Genomic Browser (Evolutionary conserved regions). Percentage similarity among species is set at 80%. Blue regions are coding exons, yellow region denotes the untranslated regions, red highlights the intergenic conservation among vertebrates. This sequence is highly conserved among human, chicken, and mouse genomes. Sequence of a peak with green star is shown in Fig 5.22B (B) ClustalW alignment of the 0.6 kb sequence. *, conserved nucleotides; putative Dlx-binding sites that contain the TAATT core in either orientation are highlighted in green, other TAAT motifs are highlighted in yellow.

<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicken</td>
<td>TTCTTAGAAAACCTAGAGCCGGAGTCGCGTTAATGCTATGGTTTGTCCATTTCAATCAGA</td>
</tr>
<tr>
<td>Human</td>
<td>-TTAAAGAAAACCTAGAGCCGGAGTCGCGTTAATGCTATGGTTTGTCCATTTCAATCAGA</td>
</tr>
<tr>
<td>Mouse</td>
<td>-TTAAAGAAAACCTAGAGCCGGAGTCGCGTTAATGCTATGGTTTGTCCATTTCAATCAGA</td>
</tr>
</tbody>
</table>

**TTT**

<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicken</td>
<td>GAGTGAGTGGCACATTCTCACCTCATTTGCAATTTTTAATTCCAATTGGACATTGCTTTT</td>
</tr>
<tr>
<td>Human</td>
<td>TAGTAAGTGGCACATTCTCACCTCATTTGCAATTTTTAATTCCAATTGGACATTGCTTTT</td>
</tr>
<tr>
<td>Mouse</td>
<td>TAGTAAGTGGCACATTCTCACCTCATTTGCAATTTTTAATTCCAATTGGACATTGCTTTT</td>
</tr>
</tbody>
</table>

**TTT**

<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicken</td>
<td>GAAACGTCCCTTCTCTTTTCTGCTGCCACTTGCAATTTTTAATTCCAATTGGACATTGCTTTT</td>
</tr>
<tr>
<td>Human</td>
<td>GAAATATCCCCTTCTCTTTTCTGCTGCCACTTGCAATTTTTAATTCCAATTGGACATTGCTTTT</td>
</tr>
<tr>
<td>Mouse</td>
<td>GAAATATCCCCTTCTCTTTTCTGCTGCCACTTGCAATTTTTAATTCCAATTGGACATTGCTTTT</td>
</tr>
</tbody>
</table>

**TTT**

<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicken</td>
<td>CGTAATGGTAGGTGCAGCTTGCTGCCACTTGCAATTTTTAATTCCAATTGGACATTGCTTTT</td>
</tr>
<tr>
<td>Human</td>
<td>TGTAATGGTAGGTGCAGCTTGCTGCCACTTGCAATTTTTAATTCCAATTGGACATTGCTTTT</td>
</tr>
<tr>
<td>Mouse</td>
<td>CGTAATGGTAGGTGCAGCTTGCTGCCACTTGCAATTTTTAATTCCAATTGGACATTGCTTTT</td>
</tr>
</tbody>
</table>

**TTT**

<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicken</td>
<td>GCAAGTGTCTTTTTATACCGCATAAAAGCCTGAAATGTTTACATATCAAAAGGCTATAGG</td>
</tr>
<tr>
<td>Human</td>
<td>GCAAGTGTCTTTTTATACCGCATAAAAGGCTGAAATGTTTACATATCAAAAGGCTATAGG</td>
</tr>
<tr>
<td>Mouse</td>
<td>GCAAGTGTCTTTTTATACCGCATAAAAGGCTGAAATGTTTACATATCAAAAGGCTATAGG</td>
</tr>
</tbody>
</table>

**TTT**

<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicken</td>
<td>ATATTGTTAGAGGGGAAGCTTTTATGCGGGGAAAAGGGAAGCTTTTATGCGGGGAAAAGGGAAGCTTTT</td>
</tr>
<tr>
<td>Human</td>
<td>ATATCGTTAGGAGGGGAAGCTTTTATGCGGGGAAAAGGGAAGCTTTTATGCGGGGAAAAGGGAAGCTTTT</td>
</tr>
<tr>
<td>Mouse</td>
<td>ATATCGTTAGGAGGGGAAGCTTTTATGCGGGGAAAAGGGAAGCTTTTATGCGGGGAAAAGGGAAGCTTTT</td>
</tr>
</tbody>
</table>

**TTT**
5.1.4 Testing the responsiveness of predicted putative enhancer elements to Dlx5 and Dlx6

We next asked whether any of these elements were responsive to Dlx proteins in a transcription assay. To this end, we co-transfected human embryonic kidney 293 cells (HEK293T) with luciferase reporter plasmids with varying amounts of a Dlx5 or Dlx6 expression plasmid.

We tested gg384, gg924, Gsc, and ggHand2 sequences. All tested sequences showed a dose-dependent response to co-expressed Dlx5 or Dlx6 proteins (Fig. 5.23). In addition, both Dlx5 and Dlx6 stimulated each reporter to similar levels.

5.2 Discussion

Enhancer prediction is one of the most intricate molecular biology problems and remains challenging for the following reasons. (1) Approximately 98% of the genome is non-coding (where enhancers are most likely to be located), (2) Enhancer sequences can be far away from the regulated genes (Hnisz et al., 2014; Kuroiwa, 2020) (3) they can be located in discrete locations and orientations which may be upstream or downstream or located in the introns of other genes (4) Computational enhancer prediction is a relatively new area and requires a large database to train the algorithms.

The VISTA Enhancer Browser is a reliable source for verifying the identification of many developmentally active enhancer elements in the mouse genome. Alone, this resource had the ability to discover putative enhancer regions with PA-specific activity in the mouse genome. Other computational resources were used alongside VISTA Enhancer Browser with the intention to improve the credibility of our discovered candidate enhancer regions. FaceBase p300 ChIP-seq captured p300-DNA associations used to predictively map developmentally active and evolutionary constrained enhancer elements. The combination of VISTA Enhancer Browser and FaceBase p300 ChIP-seq data was sufficient to identify the initial list of 470 candidate enhancer regions found within the mouse GRCm38/mm10 genome assembly.
pGL3plasmid was co-transfected with an empty expression vector and Dlx5 and Dlx6 proteins in HEK293 cells (A) corresponding VISTA putative enhancer element gg384 transcriptional activity with Dlx5 and Dlx6 proteins (B) Putative enhancer element gg924 transcriptional activity (C) putative Gsc enhancer element’s transcriptional activity (D) ggHand2 putative enhancer’s transcriptional activity with Dlx5 and Dlx6 proteins.

Fig. 5.23: Luciferase reporter assay
5.2.1 Benchmark for putative enhancer regions

The total overlap of enhancer fragments in Vista Enhancer and Face Base was 470. Seven high-quality candidate sequences were identified from the overlap with the following criteria. First, the mouse sequence must reproducibly drive expression in the PA1 of mouse embryos. Reporter expression needs to be exclusively arch-specific, and some sequences give reporter expression in other discrete tissue. Second, a sequence must be conserved among vertebrates (at a minimum between mouse, human and chicken genome), with conservation defined as greater than 80% identity between mouse and chicken sequences over a minimum of 100bp. Third, the closest protein-coding gene should be expressed in the first pharyngeal arch when jaw patterning and morphogenesis are in progress and should be within 500 kb of the CRE. Finally, since I was interested in identifying enhancers that bound Dlx protein, I required that all tested sequences contained one or more Dlx homeodomain-binding motifs.

5.2.2 In ovo testing

Of the conserved enhancer regions that were active in transgenic mouse embryos, only two orthologous chicken sequences showed any activity in the cranial neural crest cells of electroporated chick embryos. Since no embryos survived beyond 48 hr post-electroporation, any enhancer that responded to Edn signalling in post-migratory ectomesenchyme would have been missed. The advantage of using fluorescent reporter over lacZ was that, with a constitutively active mCherry expression plasmid, I could both evaluate transfection efficiency and confirm that our electroporation strategy targeted the migrating CNCCs.

Using luciferase reporter, quantification of the responsiveness of these elements to Dlx proteins showed that both Dlx5 and Dlx6 stimulated transcription through gg924 and gg384, as well as chick Hand2 and Gsc elements. This study, therefore, adds a putative enhancer region for Gsc that may be regulated by endothelin signalling and its effectors, Dlx5 and Dlx6. Furthermore, as the sequence element harbours Dlx binding sites, it suggests that Gsc may be a direct downstream target of Dlx5 in migrating CNCCs.
Chapter 6
General Discussion

6.1 Summary of work

The skull is a striking feature of the modern vertebrate. This prominent mosaic of tissues develops through the interaction of cranial neural crest cells with neighbouring tissues from all three germ layers that are the source of different signals. Signals emanating from neighbouring tissues and interpreted by NCCs regulate several gene regulatory pathways in NCCs, which play a pivotal role in craniofacial morphogenesis. Such signals include bone morphogenetic protein, fibroblast growth factor, notch, sonic hedgehog, and endothelin. These signalling pathways result in the expression of a discrete set of TFs, among them the Dlx family of regulatory proteins with demonstrated importance for patterning and morphogenesis of the articulated vertebrate jaw. This study sought to test the conservation of gene relationships between murine and avian embryos, namely the expression and endothelin-dependence of Dlx genes and their regulatory target genes.

Chapter 3 explored the expression of the Dlx gene family in pharyngeal arches, revealing that these genes in birds share a nested expression pattern during pharyngeal arch morphogenesis with strong similarities to the mouse embryo. At all stages that I examined, Dlx1 and Dlx2 were expressed in both branches of PA1 and Dlx5 and Dlx6 were expressed in md-PA1, like their orthologues in mice. Around HH22, Dlx5 transcription was also detected in mx-PA1. Dlx3 and Dlx4 transcripts were detected in a more restricted pattern in md-PA1. Of note, I showed the expression of Dlx4 for the first time in avian embryos, and it is still unclear whether Dlx4 mRNA is translated into a functional protein in the chick. Dlx4 expression is the last to appear in PA1 (around HH19) and has the most restricted expression pattern. In summary, Dlx1 and Dlx2 are expressed
throughout PA1, \textit{Dlx5} and \textit{Dlx6} are expressed throughout the mandibular prominence, and \textit{Dlx3} and \textit{Dlx4} expression is restricted medial-distally within the mandibular arch.

Next, I asked how many \textit{Dlx} genes were dependent on endothelin signalling for their PA1 expression. For this purpose, I utilized a pharmacological inhibitor of endothelin signalling, Bosentan. Bosentan antagonistically binds to Ednra and inhibits signal transduction. I specifically disrupted the Edn1 pathway around HH13 and monitored gene expression at two developmental time points; HH16-17 and HH22-23 as this corresponds to the window in which \textit{Gsc} was previously shown to be sensitive to Ednra antagonism (Kitizawa et al., 2015) in migrating NCCs. \textit{In situ} (24 hr) and qPCR (48 hr) demonstrated that \textit{Dlx3}, \textit{Dlx4}, \textit{Dlx5}, and \textit{Dlx6} are downregulated in avian pharyngeal arches after Bosentan treatment. These results show that the dependence of \textit{Dlx3} on \textit{Dlx5} and/or \textit{Dlx6}, previously demonstrated in mouse arch tissue, is likely conserved in the chick. In contrast, \textit{Dlx1} and \textit{Dlx2} expression increased when Edn signalling was reduced or inhibited.

After analyzing the Spatio-temporal expression of \textit{Dlx} genes and establishing that loss of Edn1-Ednra signalling results in the down-regulation of \textit{Dlx} homeobox genes \textit{Dlx3-Dlx6}, I asked whether chick orthologues of Dlx-responsive mouse genes were similarly regulated during pharyngeal arch patterning and morphogenesis. To this end, I reduced \textit{Dlx} gene expression with Bosentan-mediated antagonism of endothelin signalling and measured candidate gene expression. As a starting set, I chose genes with known craniofacial defects when knocked out in mice and whose expression was more than 2-fold upregulated or downregulated in md-PA1 in \textit{Dlx5/6-/-} mouse embryos (Jeong et al., 2008).

In \textbf{Chapter 4}, I selected five genes, \textit{Gsc}, \textit{Pitx1}, \textit{Hand1}, \textit{Hand2} and \textit{Gbx2} and demonstrated that all of these genes are affected by inhibiting the Edn-Dlx5/6 pathway. \textit{Gsc}, \textit{Pitx1}, \textit{Hand2}, and \textit{Hand1} showed down-regulation in the PA1 in the absence of the active signalling while the expression of \textit{Gbx2} increased. With \textit{in situ} hybridization, I also demonstrated endothelin-dependant and endothelin-independent territories within md-PA1, with Edn signalling specifically regulating the proximal domain.
Gsc was down-regulated by more than 80%, making it a promising candidate to further examine as a direct Dlx target gene. I searched the Gsc locus for conserved non-coding blocks of sequence and found a 600 bp sequence, 3.6 kb upstream of the Gsc transcription start site, that is highly conserved between humans, mice, chick, and zebrafish. The region harbours several consensus Dlx binding motifs, and this sequence behaved as a Dlx-responsive regulatory element \textit{in vitro}.

In addition to predicting novel enhancer regions in the chicken genome, I also attempted to validate chicken orthologues of some of the previously identified mouse enhancers (Chapter 5). For this purpose, I first looked for overlap between active enhancers in VISTA with ChIP-seq data from p300 binding in the mouse PA1. The binding of p300 is strongly associated with active enhancers (Heintzman et al., 2007). From this genome-wide \textit{in silico} search, we short-listed seven enhancers whose closest gene was implicated in mouse craniofacial development. I retrieved the equivalent chicken DNA sequence using a BLAT tool in the UCSC genome browser and cloned five orthologous chick sequences into ptk-GFP reporter plasmids: gg384, gg901, gg924, gg1090, and gg1114. This study also included one previously identified enhancer region from the mouse \textit{Hand2} locus (Charité et al., 2001). Only gg901 showed obvious neural crest-specific activity at 24 hr post-electroporation, and a small number of scattered GFP-positive cells were visible at 48 hr post-electroporation with the sequence cloned in the opposite orientation. The enhancer region, gg384, showed a small number of scattered GFP-positive cells at 24 hr post-electroporation in one embryo.

6.2 A proposed model of PA1 gene regulatory network in chicken

Gene regulatory networks (GRNs) are dynamic interactions between transcription factors and genomic \textit{cis}-regulatory elements (CREs). I summarized our findings and put them in the context of chicken jaw morphogenesis (Fig 6.1).

We propose a theory in which \textit{Dlx5} and/or \textit{Dlx6} are direct downstream effectors of endothelin signalling. Additionally, Dlx5 and/or Dlx6 TF also positively regulate \textit{Dlx3, Dlx4, Gsc, Pitx1, Hand1, Hand2} and \textit{Gbx2} genes in md-PA1. Furthermore, Dlx5 and/or Dlx6...
may regulate other genes like *Twist1*, and *Bmp4* and *Runx2* as their respective VISTA enhancer regions; mm901(gg901) and mm384(gg384), were active in migrating CNNC in chick embryos (see chapter 5). I cannot rule out that one or more of these genes are regulated by Dlx3 or Dlx4 in cells where those TFs are expressed.

![Diagram](image)

**Fig. 6.1: A proposed model of chick PA1 gene regulatory network**

In the PA1 model, endothelin-dependant Dlx5 is responsible for regulating and maintaining the expression of genes (chapters 3 & 4, left side) and *Twist2, Bmp4* (chapter 5, nearby enhancer regions were active in PA1). In the absence of endothelin, Dlx1, Dlx2 and Gbx2 expression is upregulated. There is also a possibility that Dlx5 and/or Dlx6 are regulating *Dlx3* and *Dlx4* and that at least some genes are directly regulated by Dlx3 and/or Dlx4. We still do not know the extent to which Dlx TFs are dependent on other factors for the regulation of their target genes. All genes shown in this figure are independently implicated in craniofacial development.

### 6.3 Is Dlx5 a master regulator in PA1?

Susumu Ohno, in 1978, coined the term “master regulator genes” (MRGs) which implies the gene functions at the top of a gene hierarchy (Kin-Chan, 2013). Recent studies modified the definition and showed that, while an MRG could be under the control of earlier-acting regulation, MRGs determine lineage or cell fate in a cell-autonomous fashion through the regulation of several downstream genes directly or indirectly and that these genes are sufficient to reprogram cell fate (Darnell et al., 1992; Kin-Chan, 2013). *Dlx5* in PA1 is regulated by endothelin signalling. *Dlx5/6* and *Ednra* knockout mice show
a homeotic transformation of the lower jaw into upper jaw structures (Beverdam et al., 2002; Clouthier et al., 2000). To see the role of Dlx5 in PA1, I utilized a gain-of-function strategy. Although we successfully misexpressed Dlx5 in mx-PA1 (Chapter 4). We did not observe the ectopic expression of any gene we tested. In conclusion, Dlx5 (and likely Dlx6) can be said to be necessary but not sufficient for mandibular fate. The inability of Dlx5 to autonomously regulate downstream genes suggests that it is not an MRG. Further experimentation in the chick embryo should include the misexpression of Edn1 in mx-PA1 to evaluate the responsiveness of candidate Dlx target genes.

6.4 Functional redundancy vs uniqueness: Do Dlx5 and Dlx6 have different transcriptional targets?

Dlx5 and Dlx6 share intragenic enhancer regions and are largely co-expressed, with Dlx5 having a somewhat broader expression than Dlx6 in chicken (this study) and mice (Jeong et al., 2008). Also, I found that Dlx5 expression extends into mx-PA1 after HH25, whereas Dlx6 does not. Jeong et al. (2008) showed that Dlx5/− and Dlx6/− mutants showed similar phenotypes, with shared common defects in craniofacial structures, but that Dlx5 null mouse embryos generally showed more severe craniofacial defects compared to Dlx6/− embryos. Furthermore, in the same experiment, Jeong et al. also demonstrated that Gbx2, Dlx4, Pou3f3, Foxl2 and Evi1/2 are more down- or upregulated in Dlx5/− null mouse embryos compared to Dlx6/− embryos. The combined Dlx5/6/− knockout also showed more robust effects on downstream genes than when either gene was knocked out alone. For example, in null Dlx5/− and Dlx6/− embryos, Gbx2 was downregulated in md-PA1, whereas its expression was completely lost in Dlx5/6/− knockout embryos (Jeong et al., 2008). This shows bigene functional redundancy and points to a compensatory mode of action at the level of individual genes as well as the jaw phenotype. There was also evidence that, in some cases, Dlx6 made a more substantial contribution to regulating some genes. For example, genes like Bmp6 and asDlx1 were more affected in Dlx6/− mouse embryos than in Dlx5/− embryos. Tmem30b gene, which is usually not expressed in mouse md-PA1, does not express in Dlx5/− mouse embryos, but it shows up-regulation in md-PA1 of Dlx6/− embryos. In a transcriptional assay, the
Gbtx2 enhancer showed two-fold more activation by Dlx6 compared with Dlx5 (Quach et al., 2016).

Also, the Hand2 enhancer was regulated by Dlx6 in PA1 (Charité et al., 2001). These findings suggest that, although each protein can compensate for the absence of the other paralogue, the proteins may be distributed to discrete target gene loci in wild-type cells.

It is still unclear how protein structure contributes to this functional redundancy. It is well known that the homeobox region of Dlx proteins interacts with a consensus binding motif to activate transcription. Though the core binding motif of all Dlx genes is the same, the flanking region also plays an important role. In general, homeobox genes need some additional flanking sequences and/or protein co-factors for binding to their consensus binding motifs. These co-factors not only give the homeoproteins increased specificity but can also alter their specificity.

6.5 Edn-Dlx5 related pathologies: Implication for human disease

Malfunction of NCCs and their derivatives can lead to several congenital disorders, including craniofacial defects. Of the described human craniofacial defects, almost 1/3 are classified as congenital disorders (Cordero et al., 2011).

In humans, at least one-third of all developmental anomalies are due to the improper development and migration of NCCs. Amongst these syndromes, many have mandible-associated craniofacial defects, including agnathia, micrognathia, retrognathia and aplasia or hyperplasia of the mandibles (Funato et al., 2016). Collectively, these are referred to as first arch syndrome (FAS) (Vieux-Rochas et al., 2010). It was also demonstrated that in FAS, the proximal structure was mainly affected as compared to distal structures (Kelberman et al., 2001; Masotti et al., 2008), which may suggest the involvement of endothelin signalling as this territory is governed by Edn-Dlx5 (Johnson et al., 2011). This needs more investigation.

Auriculocondylar syndrome (ACS) is a representative syndrome of FAS. It is an autosomal craniofacial syndrome characterized by agnathia, cleft palate,
temporomandibular joint fusion and question mark ear (Masotti et al., 2008). Studies showed that ACS is due to two mutations in Edn pathway signalling molecules, \textit{PLCB4} and \textit{GNAI3} (Rieder et al., 2012). These signalling molecules play a vital role in \textit{Dlx5} and \textit{Dlx6} regulation by Endothelin (Vieux-Rochas et al., 2010), and the level of \textit{Dlx5} and \textit{Dlx6} expression is extremely low in ACS patients (Rieder et al., 2012). Identifying factors involved in mandible specification will provide insights into the evolution of vertebrates and the mechanisms underlying congenital anomalies involving craniofacial tissues.

6.6 Future directions:

6.6.1 Effect of enhancer sequence perturbation

The most exciting future research could be the analysis of the predicted (Chapter 5) enhancer regions’ dependence on Dlx5 protein. To determine whether Dlx binding sequences are responsible for the observed gene expression, mutations of each likely Dlx-binding site in the cloned sequence can be made using a site-directed mutagenesis approach (Lee et al., 2005). The mutant amplified product can be cloned into the ptk-EGFP or pMin-luc backbones and tested \textit{in ovo} and \textit{in vitro}, respectively.

6.6.2 Determination of genome-wide Dlx5 binding sites:

In the chicken genome, putative binding sites of Dlx5 can be determined by ChIP-seq. Chromatin can be isolated from md-PA1 of 5-7 chicken embryos at HH22-24 using published protocols (Colasante et al., 2008; Zhou et al., 2004). The availability of ChIP-grade antibodies to chick Dlx5 is an issue for this approach. One solution is the infection of the PA1 with \textit{RCAS-Myc-Dlx5} or \textit{RCAS-Flag-Dlx5} virus and subsequently using a ChIP-grade antibody to the epitope tag for ChIP and library preparation. By exploiting epitope-tagging I can also compare Dlx5 and Dlx6 transcriptional targets genome-wide. Dlx5 and Dlx6 transcriptional target analysis in a given tissue will provide clues of transcriptional target redundancy and uniqueness between the two paralogues. Another extension of this approach could be the investigation of the interaction of co-activator or binding partners in PA1. Hojo et al (2016) found that Dlx5 protein is required for Sp7 anchoring on the DNA in osteoblasts. In a Chip-seq experiment, they found the AT-rich
peak highlighting the Dlx-Sp7 relationship in osteoblast specification (Hojo et al., 2016).

Another research avenue could be the structural modelling of TF binding to enhancer regions. This modelling could help in elucidating the homeobox protein interaction with its DNA binding motif. Delineating the fundamental mechanism(s) responsible for the nesting of Dlx genes expression will likely yield greater insight into vertebrate evolution.

Dlx5/6 and Edn1 receptor knockout mouse models show the same morphological phenotype. Little is known about the Dlx5/6 regulation via endothelin 1 and Mef2C pathways and their relationship. A study that can decipher the role of these regulators to activate Dlx5/6 in craniofacial patterning will shed some light on the compensatory mechanisms when one of them is knocked out.

We need to establish a loss-of-function strategy for chicken embryos to see each paralogue function. It could be either Clustered, Regularly Interspaced, Short Palindromic Repeats (CRISPR)-mediated or by small interfering RNA.

As I validated the Dlx4 gene expression, it would be interesting to know whether a functional protein is translated.

### 6.7 Significance and major contribution of research

The data produced from this study contribute to the fields of evolutionary, molecular, and developmental biology (Fig 6.2). This project focuses on transcription factors Dlx5 and Dlx6. The current study helps understand the expression of Dlx paralogues within the chick embryo and highlights the conserved expression of Dlx orthologues among the vertebrates. Comparative analysis will help in elucidating complex evolutionary processes. One of the most important attributes of this study is its insight into avian craniofacial and head development.

Neural crest cells are unique to vertebrates and have had a profound influence on vertebrate evolution. Jaw morphogenesis requires the coordination of multiple different tissues that impact GRNs in NCCs and their derivatives. Among the early-acting TFs
expressed in CNCCs, \textit{Dlx5} and \textit{Dlx6} play critical roles in lower jaw identity. I sought to understand to what extent \textit{Dlx}-responsive GRNs are conserved between mammals and birds and where differences have emerged.

![Fig. 6.2: Scope of current study](image)

The interdisciplinary work touches the boundaries of evolutionary, developmental and molecular biology.

Some original contributions of the study are summarized below:

- Documentation of the evolutionarily conserved expression of all \textit{Dlx} genes in an avian embryo, including, for the first time, the expression of \textit{Dlx4}
- Demonstration of the endothelin dependence of \textit{Dlx3} and \textit{Dlx4} in chick pharyngeal tissues
- Demonstration of the endothelin dependence of \textit{Gsc, Hand2, Hand1, and Pitx1}, thus providing at least correlative evidence for their regulation by \textit{Dlx} genes
- Demonstration that misexpression of \textit{Dlx5} in mx-PA1 is not sufficient to re-program maxillary tissue to a mandibular fate, based on a lack of ectopic marker gene expression
- Predicted candidate enhancer regions for \textit{Gsc, Twist2, Runx2, and Bmp4} in the chick genome


Colasante, G., Collombat, P., Raimondi, V., Bonanomi, D., Ferrai, C., Maira, M.,


generated four homeobox genes implicated in extra-embryonic tissue formation. *PLoS Genetics*, 10(10), e1004698. https://doi.org/10.1371/journal.pgen.1004698


