

UNIVERSITY OF OXFORD

CHD7 in CHARGE: the implications of neural crest specific regulation of chromatin remodeller CHD7 in CHARGE syndrome

---

Author:
Özlem Güneş TAYLOR

Supervisor:
Professor Tatjana SAUKA-SPENGLER

A thesis submitted for the degree of
D.Phil. in Medical Sciences

at

St. Anne's College

Trinity 2018

"You are not Dickens and this is not 'Great Expectations', although it could be "Bleak house' :) at times"

Tatjana Sauka-Spengler, 3rd October 2015

"I have been bent and broken, but - I hope - into a better shape"

Great Expectations, Charles Dickens, 1861

University of Oxford

Abstract

St. Anne's College

Trinity 2017

D.Phil. in Medical Sciences

CHD7 in CHARGE: the implications of neural crest specific regulation of chromatin remodeller CHD7 in CHARGE syndrome

by Özlem Güneş TAYLOR

The neural crest is a multipotent embryonic progenitor population. Differentiated neural crest cells contribute to a remarkable range of tissues and structures including the peripheral nervous system, bone and cartilage of the head, cardiac septum and the major vessels of the heart. CHARGE syndrome (OMIM 214800) is a sporadic, dominant disorder characterised by a multitude of congenital aberrations; ocular coloboma, heart malformations, atresia of the choanae, retardation of growth, genital hypoplasia and ear abnormalities. CHARGE syndrome is considered a neurocristopathy as the perturbed structures are significantly patterned by, or contain contributions from, the neural crest. The majority of CHARGE patients carry heterozygous mutations in chromatin remodeller CHD7. Neural crest cells lacking CHD7 have reduced expression levels of key migration markers and migrate aberrantly both *in vitro* and *in vivo*. Therefore, CHARGE syndrome could be attributed to CHD7 deficient neural crest cell migration and implicates CHD7 as a critical neural crest factor.

The neural crest has been extensively studied and a gene regulatory network detailing the hierarchical interactions between key genes compiled. Enhancers are *cis*-regulatory sites used by upstream transcription factors to promote downstream gene transcription. As such, enhancers define the hierarchical relationships between upstream transcription factor inputs and downstream gene outputs. Consequently, enhancers have been used to establish architecture of gene regulatory networks, including the neural crest gene regulatory network. To integrate CHD7 within the known neural crest gene regulatory network and establish its role within neural crest development, neural crest-specific enhancers of CHD7 and the factors utilizing them must be identified. Compellingly, 40% of CHARGE patients lack mutations within the exons of CHD7. Mutations in critical CHD7 *cis*-regulatory elements such as neural crest specific enhancers could impact CHD7 expression, accounting for these unexplained CHARGE syndrome cases.

Ten putative cranial neural crest specific *cis*-regulatory sites for CHD7 were identified by differential DNA accessibility profiling. Of the ten putative CHD7 enhancers tested by *in vivo* fluorescent reporter activity, seven demonstrated enhancing capacity and were deemed novel CHD7 enhancers. An advanced molecular toolkit was developed to further characterise these novel CHD7 enhancers, and potentially allow integration into the cranial neural crest gene regulatory network. To accommodate the limited number of cranial neural crest cells per chicken embryo, the classical neural crest model system, transcription factor biotagging was translated for small cell number chromatin immunoprecipitation. Additionally, CRISPR-Cas9 based genome- and epigenome- editing tools were generated to test the novel enhancer placement within the network.

A novel enhancer with strong cranial neural crest activity, Chd7_a, was bound by critical neural crest factors FoxD3 and Sox10. Mutations of predicted Sox10 binding sites within the novel enhancer reduced enhancing activity, as did the targeting of endogenous Sox10 by CRISPR-Cas9 genome editing. Epigenome editing of Chd7_a by repressive dCas9-KRAB activity reduced Chd7 expression levels and confirmed Chd7_a regulation of Chd7 in the chicken cranial neural crest. Functional conservation of the novel chicken Chd7 enhancer Chd7_a in humans was demonstrated by confirming enhancing activity in the equivalent CHD7_a region. Consequently, we propose chromatin remodeller Chd7 is critically regulated in the neural crest by tissue-specific transcription factors Sox10 and FoxD3 driven enhancer elements, and that perturbation of this regulation can account for the aetiology of CHARGE syndrome.

Acknowledgements

I would like to thank my supervisor Tatjana Sauka-Spengler for giving me the opportunity for this unique PhD experience. To my examiners Jim Hughes and Andrea Streit; thank you for your patient and honest assessment of this work. I would like to extend thanks to Professors Thomas Milne, Roger Patient, Richard Gibbons, and Tudor Fulga for their indispensable guidance and instrumental insight into the project. I would also like to express eternal gratitude to all the members of the Sauka-Spengler lab, past and present, both for scientific support and many happy memories. Particular thanks are due to Doctors Ruth Williams and Upeka Senanayake for their support during the experimental phase, and Doctors Amy Kenyon, Dorit Hockman, Vanessa Chong-Morrison, and Lucy Wheatley during the writing phase of this DPhil.

Thank you to those who opened the doors of their laboratories to me over countless summers: Professors Fikrettin Şahin, Timothy Aitman, Jane E Hewitt, Matt Loose, Ramiro Alberio, and John Armour.

To my scientific mentors, Professors Bob Lloyd, Conrad Nieduszynski, and Robin Lovell-Badge: thank you. You had faith when I lacked it, wisdom when I required it, and without whom none of this would have been possible.

Finally, I would like to thank my incredible friends and family for their unerring belief, support, and encouragement along the way.

PRB, ESW, KS, KHJ, EJH, AP, ACM, MH, VW, DD, SET, BAKT, AT.

<3

Contents

Abstract **v**

Acknowledgements **vii**

List of Figures **xv**

List of Abbreviations **xvii**

1 Introduction **1**

1.1 The Neural Crest 1

1.2 NC Axial Sub-populations & Derivatives 7

1.2.1 Cranial NC 7

1.2.2 Cardiac and Vagal NC 9

1.2.3 Trunk and Sacral NC 10

1.3 The Neural Crest Gene Regulatory Network 11

1.3.1 Enhancers provide spatio-temporal control and modularity . 13

1.3.2 Chromatin remodellers pattern global chromatin landscapes . 15

1.4 CHARGE syndrome: a Neurocristopathy 18

1.4.1 Neurocristopathies are characterised by aberrant NC development 18

1.4.2 CHARGE syndrome 18

1.4.3 Mutations in chromatin remodeller CHD7 cause CHARGE syndrome 20

1.5 Fitting Chromatin Remodeller CHD7 into the Cranial NC-GRN . . . 24

1.5.1 Identifying enhancers using the classic chicken embryo model system . . . 26
1.5.2 Transcription Factor binding site profiling in chicken . . . 27
1.5.3 Genome- and epi-genome editing technologies in chicken . . 27
1.6 Project Summary & Objectives . . . 28

2 Materials & Methods **33**
2.1 Chicken Embryo Methods . . . 33
2.1.1 Chicken embryo incubation . . . 33
2.1.2 Chicken embryo collection and electroporation . . . 33
2.2 Putative *Cis*-Regulatory Element Selection and Enhancer Testing Methods . . . 34
2.2.1 Identification of putative *cis*-regulatory elements . . . 34
2.2.2 Functional testing of putative enhancers . . . 35
2.3 Whole-Mount Imaging Methods . . . 36
2.3.1 Labelled RNA probe generation . . . 36
2.3.2 Wholemount *in situ* hybridisation . . . 37
2.3.3 Tyramide fluorescent *in situ* hybridization . . . 38
2.3.4 Hybridisation Chain Reaction (HCR) fluorescent *in situ* . . . 39
2.3.5 Whole-mount immunocytochemistry . . . 39
2.4 Molecular Methods . . . 40
2.4.1 RNA extractions . . . 40
2.4.2 cDNA generation . . . 40
2.4.3 RT-QPCR analysis . . . 40
2.4.4 NanoString gene expression profiling . . . 41
2.4.5 Biotagging-compatible construct cloning . . . 41
2.4.6 Avi-FoxD3-EGFP construct cloning . . . 42
2.4.7 Citrine-FoxD3-Avi construct cloning . . . 42
2.4.8 FoxD3-Avi construct cloning . . . 43

2.4.9 Avi-Sox10 construct cloning 43
2.4.10 Biotin-ChIP-QPCR 43
2.5 Genome- and Epigenome-Editing Toolkit Methods 44
2.5.1 sgRNA cloning 44
2.5.2 pX330-Cas9 construct 45
2.5.3 pX330-dCas9-KRAB construct cloning 45
2.5.4 dCas9-KRAB construct cloning 45
2.5.5 pCI-dCas9-KRAB-GFP 45
2.5.6 High Resolution Melt Analysis (HRMA) analysis 46
2.5.7 T7 Endonuclease Assay 46
2.6 *In vitro* Protein Methods 47
2.6.1 Protein extraction from chicken embryos 47
2.6.2 Western Blotting 48

3 Establishing an Advanced Toolkit for Exploring Gene Regulatory Interactions using the Chicken Embryo Model System 49
3.1 Introduction 49
3.1.1 Resolving Gene Regulatory Networks 51
3.1.2 Tools currently available for studying enhancers 51
3.1.3 The chicken model system 53
3.1.4 Using the chicken system to resolve the NC-GRN 54
3.1.5 Current tools used to dissect the NC-GRN in chicken embryos 55
3.1.6 Expanding the chicken toolkit 57
3.2 The 'Biotagging' Principle and Toolkit for Chicken 58
3.2.1 Biotagging key NC TFs FoxD3 and Sox10 for use in the chicken model system 59
3.2.2 Characterising Avi-tagged TF constructs expression and confirmation of Avi-tagged FoxD3 protein *in vivo* biotinylation . . 61

3.2.3 Immunocytochemistry shows differences between N- and C-terminally Avi-tagged FoxD3 proteins *in vivo* 63
3.2.4 Streptavidin accessibility of Avi- tagged FoxD3 proteins confirmed *in vivo* 68
3.2.5 Nanostring nCounter analysis showed no major perturbations of key NC genes by exogenous expression of Avi-tagged FoxD3 68
3.2.6 ChIP-QPCR 70
3.3 CRISPR-Cas9 Mediated TF Knockdown in Chicken 71
3.3.1 TF knockdown in chicken via CRISPR/Cas9 mediated Non-Homologous End Joining (NHEJ) 72
3.3.2 Constructs generated for CRISPR-Cas9 application in chicken 73
3.3.3 CRISPR-dCas9 based epigenome engineering 74
3.3.4 Epigenome engineering via targeted demethylation/compaction of chromatin surrounding enhancers 75
3.4 Summary of Advanced Toolkit Developed to Explore Gene Regulatory Networks in the Chicken 77

4 Identifying Putative Cranial NC Specific *Cis*-Regulatory Elements for Chromatin Remodeller Chd7 in the Chicken Embryo 79
4.1 Introduction 79
4.2 Results 89
4.2.1 Onset of Chd7 expression overlaps with key NC TFs FoxD3 and Sox10 in the chicken embryo 89
4.2.2 Comparative chromatin accessibility profiling indicates multiple putative *cis*-regulatory regions within the epigenomic landscape surrounding Chd7 in cranial NCCs. 90
4.2.3 Seven enhancers identified in a screening of putative cranial NC specific Chd7 *cis*-regulatory regions 92
4.3 Discussion 96

5 Placing a Novel NC-specific Chd7 Enhancer Element within the NC-GRN 101

5.1 Introduction 101

5.2 Results 103

5.2.1 Chd7_a enhancer activity overlaps with that of Sox10 enhancer Sox10E2 103

5.2.2 A highly accessible 850bp region contains the essential regulatory core of the Chd7_a enhancer element 103

5.2.3 Avi-FoxD3 and Avi-Sox10 bind novel enhancer Chd7_a 105

5.2.4 Sox10 binding is essential for Chd7_a enhancing activity . . . 107

5.2.5 Targeted epigenomic modification of Chd7_a chromatin context affects Chd7 expression levels 111

5.2.6 Chd7_a is syntenically conserved in human 113

5.3 Discussion 116

6 Conclusion and Future Directions 119

A Putative Chd7 enhancers 127

B QPCR primer details 131

C CRISPR toolkit component primers 133

D Nanostring Codeset accession numbers 135

E Chicken Chd7 *in situ* probe details 139

F Annotation of Chd7_a enhancer core region predicted sites 141

List of Figures

1.1 Stages of chicken NC ontogeny 3
1.2 NC axial subpopulations and respective derivatives 8
1.3 NC-GRN 14
1.4 Cranial NC expression of FoxD3 is faithfully recapitulated by NC1 enhancer 16
1.5 Clinical presentation of CHARGE syndrome patients 21
1.6 Schematic of CHD7 protein and functional domains 24
1.7 CRISPR-Cas9 technology schematic 29

3.1 Chicken embryo *ex-ovo* electroporation 56
3.2 Biotagging and biotin-ChIP-sequencing 60
3.3 Chicken biotagging constructs 62
3.4 Western Blot comparison of all Avi-tagged FoxD3 fusion proteins . . 63
3.5 FLAG-tag in undetectable in Citrine-FoxD3-Avi as compared to Avi-FoxD3-EGFP 65
3.6 Removal of Citrine reporter on C-terminally Avi-tagged FoxD3 releases FLAG detection *in vivo* 66
3.7 Comparison of Avi-FoxD3 and FoxD3-Avi expression *in vivo* 67
3.8 Immunocytochemistry to confirm streptavidin *in vivo* accessibility to biotinylated Avi-tagged FoxD3 protein 69
3.9 Analysis of the effect of ectopic expression of Avi-tagged FoxD3 protein using Nanostring technology 70
3.10 CRISPR-Cas9 genome engineering constructs 75

3.11 Determining U6 promoter choice for chicken CRISPR vectors using gRNA to FoxD3 76
3.12 dCas9-KRAB targeting of active enhancers 77

4.1 Expression of FoxD3, Sox10 and Chd7 in the cranial region of chicken embryos at HH stage 8 to 11 91
4.2 The genomic landscape surrounding the Chd7 locus in the chicken . 92
4.3 Table of putative Chd7 NC enhancers 93
4.4 Positive Chd7 enhancers 95

5.1 Chd7_a enhancer activity overlaps with Sox10E2 enhancer expression 104
5.2 Isolation of the core region of the Chd7_a enhancer 105
5.3 Analysis of putative TF binding sites within the Chd7_a enhancer . . 108
5.4 Avi-Sox10 and Avi-FoxD3 Biotin-ChIP-QPCR 110
5.5 CRISPR-Cas9 editing of the Sox10 locus results in loss of Chd7_a activity 112
5.6 Targeting of Chd7_a using dCas9-KRAB results in down-regulation of Chd7 gene expression 114
5.7 Syntenic human CHD7_a enhancer is active *in vivo* in chicken embryos 115

List of Abbreviations

ATAC	Assay for Transposase-Accessible Chromatin
CHARGE	Coloboma, Heart malformations, Atresia of the choanae, Retardation of growth, Genital hypoplasia and Ear abnormalities syndrome
ChIP	Chromatin-Immunoprecipitation
dCas9	catalytically-dead Cas9
EMT	Epithelial-to-Mesenchymal Transition
FACS	Fluorescence Activated Cell-Sorting
GRN	Gene Regulatory Network
hESCs	human Embryonic Stem Cells
HH	Hamburger Hamilton stage
HRMA	High Resolution Melt Analysis
ICF	Immunodeficiency, Centromere instability, Facial anomalies syndrome
NC	Neural Crest
NC-GRN	Neural Crest Gene Regulatory Network
NCC	Neural Crest Cell
NCCs	Neural Crest Cells
NGS	Next Generation Sequencing
NHEJ	Non-Homologous End Joining
NLS	Nuclear Localisation Signal
NT	Neural Tube
PCR	Polymerase Chain Reaction
PGCs	Primordial Germ Cells
QPCR	Quantitative Polymerase Chain Reaction

RNAi	RNA interference
RT-QPCR	Real Time quantitative Polymerase Chain Reaction
seq	sequencing
sgRNA	single guide RNA
shRNA	short hairpin RNA
ss	somite stage
TAD	Topologically Associated Domain
TF	Transcription Factor
TFs	Transcription Factors

Chapter 1

Introduction

1.1 The Neural Crest

The neural crest (NC) is an ectodermally-derived, transient progenitor population unique to vertebrates. Initially residing along the entire dorsal aspect of the neural tube (NT), neural crest cells (NCCs) migrate away to distal portions of the embryo. An obligatory epithelial-to-mesenchymal transition (EMT) allows the ectodermally-derived NCCs to migrate from the NT thereby facilitating their ultimate contributions to a plethora of unrelated tissues. Remarkably, NCCs contribute to mesenchymal structures including the thyroid and cardiac septal tissue, bone and cartilage of the craniofacial skeleton (reviewed by Bronner and LeDouarin (2012)) (Hall, 2000; Dupin et al., 2006; Douarin and Kalcheim, 1999; Adams and Bronner-Fraser, 2009; Creazzo et al., 1998). Other NC derivatives include neuronal, secretory, and pigment cell-types e.g. neurons and glia of the peripheral nervous system, adrenal chromaffin cells and melanocytes (reviewed by Le Douarin and Kalcheim (1999) (Bhatt et al., 2013; Sommer, 2011)). Consequently, NC dysregulation leads to varied and clinically significant congenital irregularities or outcomes. Human syndromes characterised by NC-derived structural aberrations are collectively called neurocristopathies (Bolande, 2010). To establish the causes of neurocristopathies, and ultimately harness the therapeutic potential of NCCs, the molecular mechanisms underlying normal NC development must first be understood.

Classically, NC ontogeny is divided into a number of characteristic stages; induction, specification, migration, and differentiation. This classical model postulates that NCCs are induced at the border between the neural- and non-neural ectoderm during neurulation (reviewed by Baker and Bronner-Fraser (1997) and Milet and Monsoro-Burq (2012) (Liem et al., 1995; García-Castro et al., 2002; Selleck and Bronner-Fraser, 1995; Moury and Jacobson, 1989)). Neurulation is the gradual separation of neural plate from non-neural ectoderm during early embryogenesis. This ultimately shapes the neural plate into the NT beneath the non-neural ectoderm, and the specified NC comes to reside at the dorsal aspect of the newly-formed NT (see Figure 1.1.). While the NT forms the central nervous system, the remaining non-neural ectoderm eventually covers the entire embryo as the epidermis. NCCs then undergo the aforementioned EMT, migrate from the NT, and ultimately differentiate at their final destinations (reviewed by Sauka-Spengler and Bronner-Fraser (2008)). Supporting this classical induction model, juxtapositions of epidermis and neural plate tissues in amphibian embryos result in the appearance of NC like-cells at these artificially generated junctions (Moury and Jacobson, 1989; Selleck and Bronner-Fraser, 1995; Kalcheim and Sela-Donenfeld, 2000). Furthermore, transplantation of quail neural plate tissue into chicken non-neural ectoderm results in NCCs from both donor and host tissues (Selleck and Bronner-Fraser, 1995). The importance of mesodermal signalling in NC induction has also been highlighted by grafting experiments (reviewed by LaBonne and Bronner-Fraser (1998)). Transplantation of paraxial mesoderm to neural plates in the chicken, or the dorsolateral marginal zone to animal caps in Xenopus, led to an apparent induction of NCCs (Monsoro-Burq et al., 2003; Kvon, 1945; Bonstein et al., 1998; Marchant et al., 1998). Thus, NC induction is attributed to a convergence of signals emanating from the mesoderm, neural- and non-neural ectoderm upon the neural plate border region (reviewed by Baker and Bronner-Fraser (1997) (Steventon et al., 2009)).

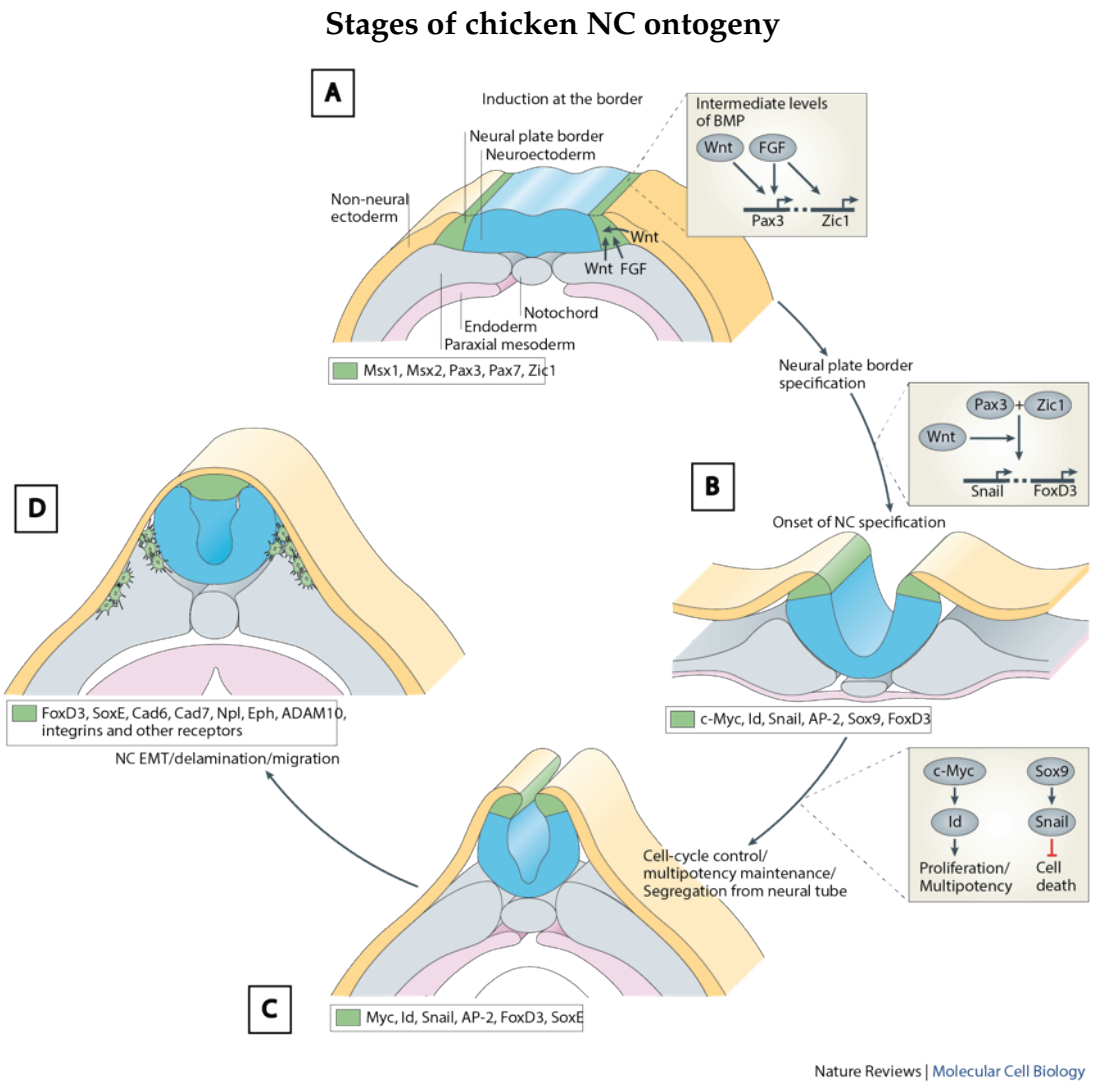


FIGURE 1.1: (A) NC development begins at the gastrula stage, at the edges of the neural plate in the neural plate border. (B) As the NT is formed from the neural plate, NC progenitors are specified at the dorsal aspect of the neural folds. (C) Subsequent to specification, NCCs undergo EMT and initiate delamination from the NT. (D) NCCs migrate along stereotypical pathways to varied destinations, where they will eventually give rise to diverse derivatives. (F) List of diverse NC derived cell types, including cells of mesenchymal, neuronal, secretory and pigmented identity. (Image from Sauka-Spengler and Bronner-Fraser (2008))

However, the traditional model of fate restriction over developmental time stipulates a progressive loss of pluripotency from stem cell to differentiated cell, via commitment to broad lineages e.g. ectoderm, mesoderm, endoderm. Cells committed to the ectoderm lineage are theoretically limited to ectodermal derivatives. Thus, the contribution of mesenchymal cell types by the ectodermally-derived NC indicates a stem cell-like pluripotency (Bronner-Fraser and Fraser, 1988; Baggiolini et al., 2015). Therefore, this classical NC induction model necessitates specified ectoderm cells within the neural plate border to effectively re-acquire multipotency in establishing their NC fate. An alternative explanation of the surprising stem-cell-like ability of NCCs to contribute mesenchymal cell types is that NCCs are not induced, but rather maintain their pluripotency from the stem-cell stage and are intermingled with the surrounding ectodermal cell types . Early experiments transplanting tissues from differentially pigmented Urodeles indicated a shared lineage between NCCs, the epidermis and the neural plate (Moury and Jacobson, 1989). Further evidence of a stem-cell like multipotency of the NC comes from experiments using chicken embryos which showed individually labelled cells from the dorsal NT could contribute both NC and NT derivatives (Bronner-Fraser and Fraser, 1988, 1991, 1989; Frank and Sanes, 1991). Other work has demonstrated that regions of the chicken epiblast expressing Pax7 are able to generate NCCs in the absence of exogenous signals under neural explant conditions containing future neural and non-neural ectoderm only. Despite the absence of mesodermal signalling, these Pax7 epiblast explants were able to differentiate into melanocytes *ex-vivo*, an exclusively NC derivative (Basch et al., 2006). Contrary to the established induction model, these findings suggests that NCCs are specified prior to neurulation, and possibly do not require mesodermal signalling to acquire their characteristic multipotency.

Further contradicting the classical model of signalling induced NC multipotency,

in 2015, Buitrago-Delgado *et al* presented striking evidence supporting the hypothesis that the NC retains its pluripotency from the stem-cell like blastula stage (Buitrago-Delgado et al., 2015). Encouraged by a growing number of reports detailing emerging roles for NC factors Sox5, Myc and their transcriptional targets Id3, FoxD3 and Snail1 in embryonic stem cells (Bellmeyer et al., 2003; Cartwright et al., 2005; Liu and Labosky, 2008; Lin et al., 2014) Buitrago-Delgado *et al* found that in addition to requiring the core pluripotency network factors Pou5F1 (Oct91, 25 and 60 in Xenopus), Sox2 and Nanog (Ventx1/2 in Xenopus), blastula stage Xenopus cells co-expressed classical NC transcription factors (TFs) FoxD3, Snail1, Id3, TF-AP2 and Ets. As such, Buitrago-Delgado *et al* redefined the neural plate border factors as general pluripotency maintenance factors. However, recent work in zebrafish comparing the expression of genes in single FoxD3 positive cells at 50% epiboly and 5-6ss pre- migratory NCCs, showed that FoxD3 positive NC cells maintain their pluripotency using different sets of paralogous factors. As indicated by Buitrago-Delgado *et al*, the majority of epiboly stage FoxD3 positive cells expressed orthologs to Oct4, SoxB, Nanog and Vox. However, the pre-migratory stage FoxD3 positive NCCs did not express this combination of genes, instead they suggest an alternative set of paralogous factors are used to maintain NCC pluripotency (Lukoseviciute et al., 2018). This challenges the proposition that NCCs are residual blastula stage cells (Buitrago-Delgado et al., 2015) but rather suggests a paralogous gene regulatory network driving multipotency specifically in the NC.

NCCs are morphologically indistinct from the neighbouring NT-forming neuroepithelial cells during neurulation. NCCs initially express cell adhesion molecules, such as N- cadherin, which serve to maintain tight cell junctions in epithelial contexts. However, NCCs undergo an obligatory epithelial- to-mesenchymal transition (EMT) to allow their delamination from the NT. During EMT, NCCs are converted from a tightly adherent epithelial layer into a more dispersed mesenchymal phenotype. This physical separation of NCCs from their neighbouring cells leaves them morphologically intact and allows their gradual expulsion from the NT epithelium

in a process called delamination. NC delamination is necessary to allow NCC migration from the NT. During EMT NCCs upregulate the expression of factors which dissociate tight cell junctions (Savagner et al., 1997) and downregulate N-cadherin until they reach their eventual post- migratory re-aggregation at sites such as the sympathetic and dorsal root ganglia (Akitaya and Bronner-Fraser, 1992). E-cadherin is necessary for NC migration *in vivo*. Morpholino knock- down of E-cadherin cannot be rescued by co-injection of other cadherins in *Xenopus laevis* (Huang et al., 2016).

Interestingly, in addition to the acquisition of mesenchymal traits, EMT is thought to impart stem-cell properties such as expression of stem-cell markers and acquisition of self-renewal capacity to the transitioning cell population (Chang et al., 2011; Mani et al., 2008; Morel et al., 2008; Ahlstrom and Erickson, 2009; Strobl-Mazzulla and Bronner, 2012). This raises the intriguing possibility that the onset of migration and NC multipotency are linked and might be simultaneously controlled as part of the EMT program. Intriguingly, enforced expression of key EMT TFs can turn mammary epithelial cells into stem-like cells thus causing breast cancer cells to adopt a tumour-initiating phenotype (Scheel and Weinberg, 2011).

Finally, it is worth noting that NCC migration does not occur in isolation, but instead is dependent on the movement of another precursor population- the placodes. Placodes are embryonic structures, originally ectodermal thickenings, formed of placodal cells that eventually give rise to sensory organs of the cranial region (reviewed by Streit (2004); Graham and Shimeld (2013)). Initially, placodal cells are intermingled with NCCs in the neural plate border region. NCCs are hypothesised to 'chase' placodal cells via Sdf1 based chemotaxis, while PCP and N-cadherin signalling causes the 'run' repulsion of the placodal cells. These interactions are essential for fidelity of NC migration and for segregation of the placodes (Theveneau et al., 2013).

1.2 NC Axial Sub-populations & Derivatives

NCCs reside along the entire anterio-posterior axis of the vertebrate embryo and are divided into five overlapping, but functionally distinct, axial subdivisions; cranial, cardiac, vagal, trunk and sacral (Sauka-Spengler and Bronner, 2010; Kelsh and Erickson, 2001; Zhang et al., 2014). NC- EMT permits NCC migration along stereotypic pathways from the NT, mingling with other mesenchymal cell types prior to differentiation. At different axial levels, NCCs commence their migration differentially and contribute different derivatives (see Figure 1.2.). For example, all levels contribute melanocytes, however only the cranial region contributes the craniofacial derivatives.

1.2.1 Cranial NC

Cranial NCCs migrate dorso-laterally from the NT and produce the cranial neurons, glia and cartilage, bone and connective tissues of the face (See Figure 1.1. & 1.2.). Vertebrate cranial sensory ganglia contain both NC and placode derived neurons. Specifically, the proximal sensory neurons of the trigeminal (Vth), facial (VIIth), glossopharyngeal (IXth), and vagal (Xth) cranial ganglia, in addition to peripheral glia, are NCC derived (reviewed by Barlow (2002)) (Le Douarin and Kalcheim, 1999).

Additionally, cranial NCCs that enter the pharyngeal arches and pouches, give rise to the bones of the middle ear and jaw, odontoblasts and thymic cells. Pharyngeal pouches form between the pharyngeal arches, outgrowths of the embryonic head and neck that the NCCs migrate into, and eventually become the thymus, thyroid and parathyroid. Chicken cranial NCCs migrate in three major canonical pathways from the area anterior to rhombomere 6 (Lumsden et al., 1991). NCCs from rhombomeres 1 and 2 migrate via the first pharyngeal, or mandibular, arch to form the jaws and incus- and malleus- bones of the inner ear. NCCs from the 4th rhombomere populate the second pharyngeal arch and form the hyoid cartilage of

NC axial subpopulations and respective derivatives

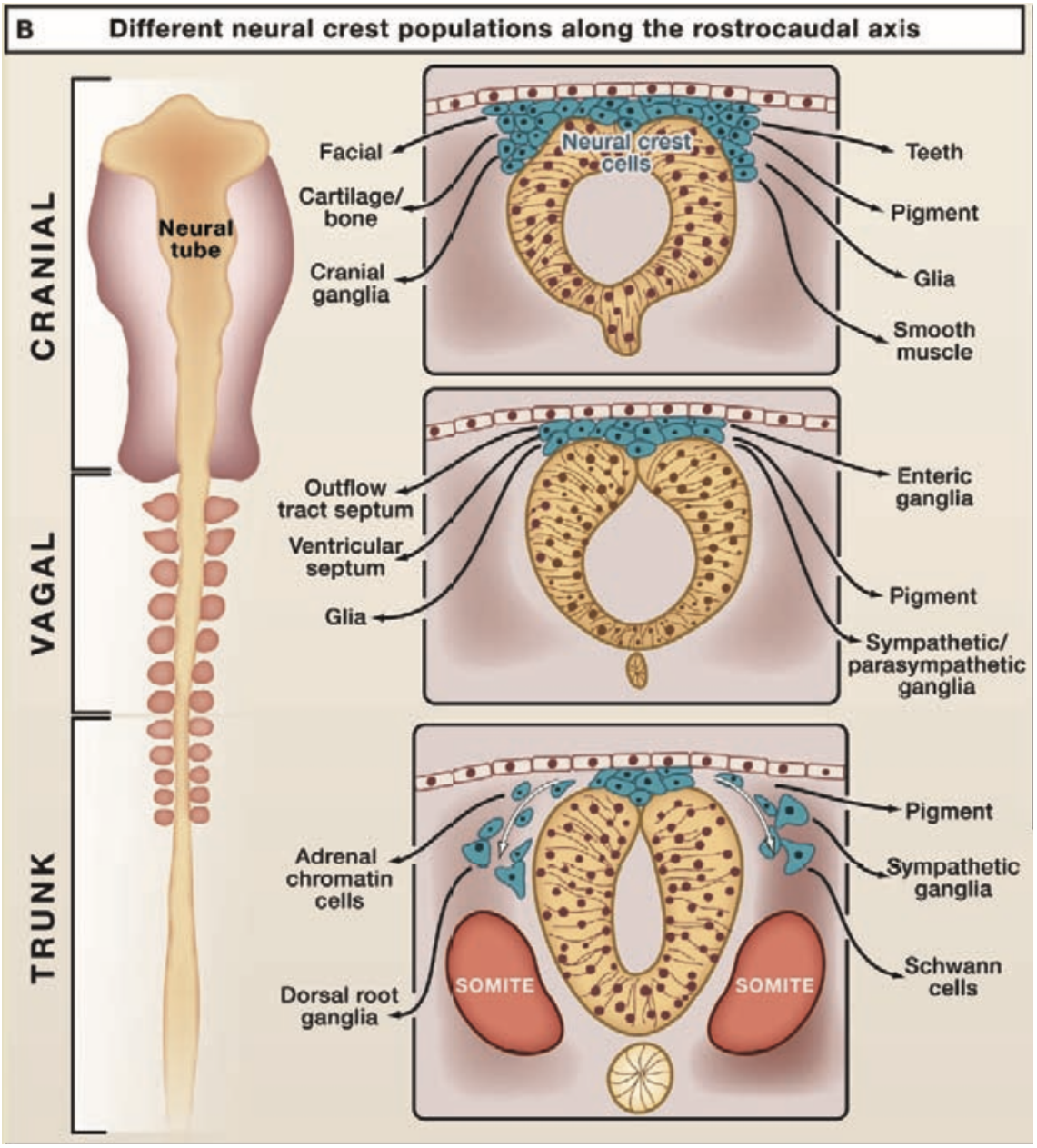


FIGURE 1.2: Diagram of different NC subpopulations along the rostro-caudal axis of the chicken embryo and their respective derivatives and embryonic contributions. Cardiac NC and sacral NC are not marked. (Image from (Sauka-Spengler and Bronner, 2010))

the neck. NCCs from the 6th rhombomere form the thymus, parathyroid and thyroid glands (Bockman and Kirby, 1984). Only the cranial NC is able to differentiate into cartilage and bone; indicating the cranial region retains the most multipotent potential of all the axial levels. Consequently, the cranial NC is of major clinical significance.

1.2.2 Cardiac and Vagal NC

The cardiac NC is located between the anterior cranial NC and the more posterior trunk NC, from the posterior border of somite three up to the mid-otic placode at its anterior (See Figure 1.2.). Despite the cardiac NC overlapping with the anterior portion of the vagal NC, NCCs from other axial levels are unable to compensate for the removal of cardiac NC in chicken (Kirby, 1987; Kirby et al., 1987; Kirby and Waldo, 1990). Cardiac NCCs migrate ventrally to the circum-pharyngeal ridge, before moving into the pharyngeal arches. NCCs that migrate beyond the third, fourth and sixth branchial arches, entering the aortic sac and cardiac outflow tract, participate in the formation of the musculo-connective tissue wall of the arteries as they form from the heart. They additionally contribute to the septum parting pulmonary circulation from the aorta (Le Lievre and Le Douarin, 1975). Interestingly, mouse mutants lacking the TF Pax3 in the cardiac NC present with a failure to separate the aorta and pulmonary artery, in addition to defects of the thyroid, parathyroid and thymic glands. This is significant, as congenital heart defects in mice and humans often co-present with defects in these glands; implicating the NC (Conway et al., 1997).

Vagal NCCs reside along the NT between somites 1-7 of the chicken embryo and colonise the embryonic gut where they generate parasympathetic ganglia and form a key part of the enteric nervous system (See Figure 1.2.) (Le Douarin, 1973; Pomeranz et al., 1991). The sacral NC also contributes to this enteric nervous system but lies posterior to somite 28 of the chicken embryo and as such only contributes to the

colonic innervation. Vagal NCCs innervate the entire remaining tract. Mutations affecting the vagal and sacral NC are implicated in diseases such as Hirschsprung's disease. Hirschsprung's congenital disorder characterised by a lack of gastrointestinal innervation causing a build-up of stool and often presenting soon after birth. The severity of the Hirschsprung's disease phenotype reflects the amount of gut lacking innervation, which is thought to occur as a consequence of faulty NCCs migration and a loss of reaching their final destinations (Arshad et al., 2012).

1.2.3 Trunk and Sacral NC

The trunk contributes three classes of derivatives: melanocyte, neurons and glia (See Figure 1.2.) (Weston, 1991; Henion and Weston, 1997; Raible et al., 1992). Trunk NCCs acquire their final fates as they migrate away from the NT via two distinct canonical pathways. These streams maintain their integrity via Eph patterning, as shown by the mixing of migrating NCC streams in Xenopus upon blocking of Eph receptors (Smith et al., 1997). NCCs acquiring melanocyte fate migrate dorsolaterally, entering the ectoderm though the basal lamina, and travel through the dermis until they reach the ventral midline. Transplantation of pigmented donor chicken NT tissue into albino host strains demonstrated this path or migration (Rawles, 1948). The second migration pathway navigates via the sclerotome, an embryonic somite which contributes bone and other skeletal components such as cartilage (Ordahl and Le Douarin, 1992). NCCs which migrate ventro- laterally via the anterior portion of the sclerotome eventually form sympathetic ganglia, aortic nerve clusters and the adrenal medulla. Trunk NCCs in the chicken adrenal medulla are induced to become sympathetic neurons in the presence of paracrine factors (Varley et al., 1995). NCCs which remain within the sclerotome form sensory neurons containing dorsal root ganglia (Frank and Sanes, 1991). Sacral NCCs contribute to the enteric nervous system, pelvic ganglia and innervation of the lower urinary tract (Wiese et al., 2017).

Whilst it remains to be seen if NCCs are still pluripotent when they migrate away from the crest of the NT (reviewed by Le Douarin and Dupin (2003)), the environment into which NCCs migrate largely appears to define their final differentiation state. For example, leukaemia inhibition factor (LIF) has been shown to encourage the trans- differentiation of NC-derived adrenergic sympathetic neurons into cholinergic neurons without affecting their overall viability (Yamamori et al., 1989; Fukada, 1980). However, it is also clear that many of the factors involved in NC differentiation are used reiteratively but that they have context dependent roles (Anderson, 1989). For example, endothelin-3 stimulates NCC differentiation into adrenergic neurons in the gut but melanocyte differentiation in the skin (Baynash et al., 1994; Lahav et al., 1996). Taken together, NCCs are a complicated and diverse group of cells which appear to be primed towards different fates prior to disembarking from the NT but apparently also retain some characteristic stem-cell-like plasticity even subsequent to their differentiation (further reviewed by Kelsh and Erickson (2001); Dupin and Sommer (2012)).

1.3 The Neural Crest Gene Regulatory Network

Starting from a single cell, embryonic development proceeds in an orderly and irreversible manner to the production of a whole animal. This process is rigorously governed by the coordinated actions of large sets of genes, expressed in precise spatio-temporal patterns (Emmert-Streib et al., 2014a; Davidson and Levine, 2005). Gene regulatory networks (GRNs) are summaries of the many complex interactions underlying, and orchestrating, spatio-temporally regulated processes during development. This precise spatio-temporal control is achieved by the combined activity of transcriptional regulator inputs, including TFs, upon *cis*-regulatory sequences such as enhancers, silencers and insulators (Davidson et al., 2003; Davidson and Levine, 2008). GRNs detail whether the interactions are known to be direct or indirect, and whether they require the coordinated actions of many regulators (Davidson et al.,

2003). GRNs describe the effect on downstream gene expression by the regulation exerted by upstream regulators via these *cis*-regulatory sequences (Davidson and Levine, 2005, 2008). As such, GRNs define the specific links within developmental processes, and characterise the interactions between the regulatory elements underlying them (Shlyueva et al., 2014; Emmert-Streib et al., 2014a).

Elucidating the regulatory interactions defined by GRNs is crucial to our improved understanding of the molecular mechanisms underlying human development and disease. It has been suggested that the information within GRNs could be used as biomarkers for diagnostic, predictive and prognostic purposes in a clinical setting (Ben-Hamo and Efroni, 2011; Matthias Dehmer and Emmert-Streib, 2013). GRNs have also been used for more efficient design of drugs, hold potential for patient-specific induced pluripotent stem cell therapies and will be significant in the development of personalised medicines (Emmert-Streib et al., 2014a,b). However, this would require a deep understanding of the GRNs in question, and more significantly of how different GRNs interact with each other within an organism. While machine learning techniques and technologies are vastly improving and would not constitute a rate-limiting-step, the experiments necessary to determine this level of knowledge do pose a significant challenge. As such, this level of rational design is still some time away.

Due to their characteristic stem cell-like multipotency, and implications in human disorders, there is increasing interest in using NC cells for therapeutic purposes. Therefore, elucidating the GRN underlying the development of this complex, and clinically important, progenitor population is essential. Since its first description (His, 1868). NC development has been intensively investigated in amphibian, avian and mammalian embryos. These investigations into the regulation of NC development, and the differences between the different axial levels, has led to the generation a complex gene regulatory network (GRN) which describes how different gene products from TFs to cell adhesion molecules interact to guide the normal process of NC development (see Figure 1.3.) (Sauka-Spengler and Bronner-Fraser,

2008). The NC-GRN is a tool for understanding the stem-cell like properties of the NC, and a framework within which to better hypothesise about how different neurocristopathies might come about. Specifically, there are different sub-networks for the different axial levels, as their spatio- temporal regulation is evidently different on account of differential migration (Meulemans and Bronner-Fraser, 2004; Sauka-Spengler and Bronner-Fraser, 2008). Diagram of the current NC-GRN detailing the known gene interactions at the neural plate border, pre-migratory, and migratory stages. Additionally, known network elements from derivative populations such as neurons/glia, melanocytes and chondrocytes are also shown here. Enhancer details are not shown in this diagram. Signalling inputs from the neural plate, epidermis and mesoderm at the neural plate border stage are also shown here. All abbreviations used within this diagram are detailed to the right of the network. (Image from (Betancur et al., 2010b))

1.3.1 Enhancers provide spatio-temporal control and modularity

Enhancers are *cis*-regulatory elements that facilitate upstream TF mediated recruitment of RNA polymerase II to core promoters/ proximal regulatory elements at transcription start sites. This recruitment serves to promote downstream gene transcription leading to increased gene products (Khan et al., 2013). As such, enhancers are used to establish architecture of these GRNs by defining the relationships between upstream TF inputs and downstream gene outputs (Davidson and Levine, 2008). Active enhancers occupy regions of open chromatin, making them DNase-sensitive. While H3K27Ac and H3K4me1 also mark promotors, enhancers are typically flanked by a combination of H3K4me1, H3K4me2 and H3K27Ac histone modifications. Enhancers are also reported to be bound by acetyl-transferases like p300, by Mediator subunits, chromodomain helicases, cohesin binding factors and RNA polymerase II (Kolovos et al., 2012; Shlyueva et al., 2014; Emmert-Streib et al., 2014a; Ong and Corces, 2012; Whalen et al., 2016).

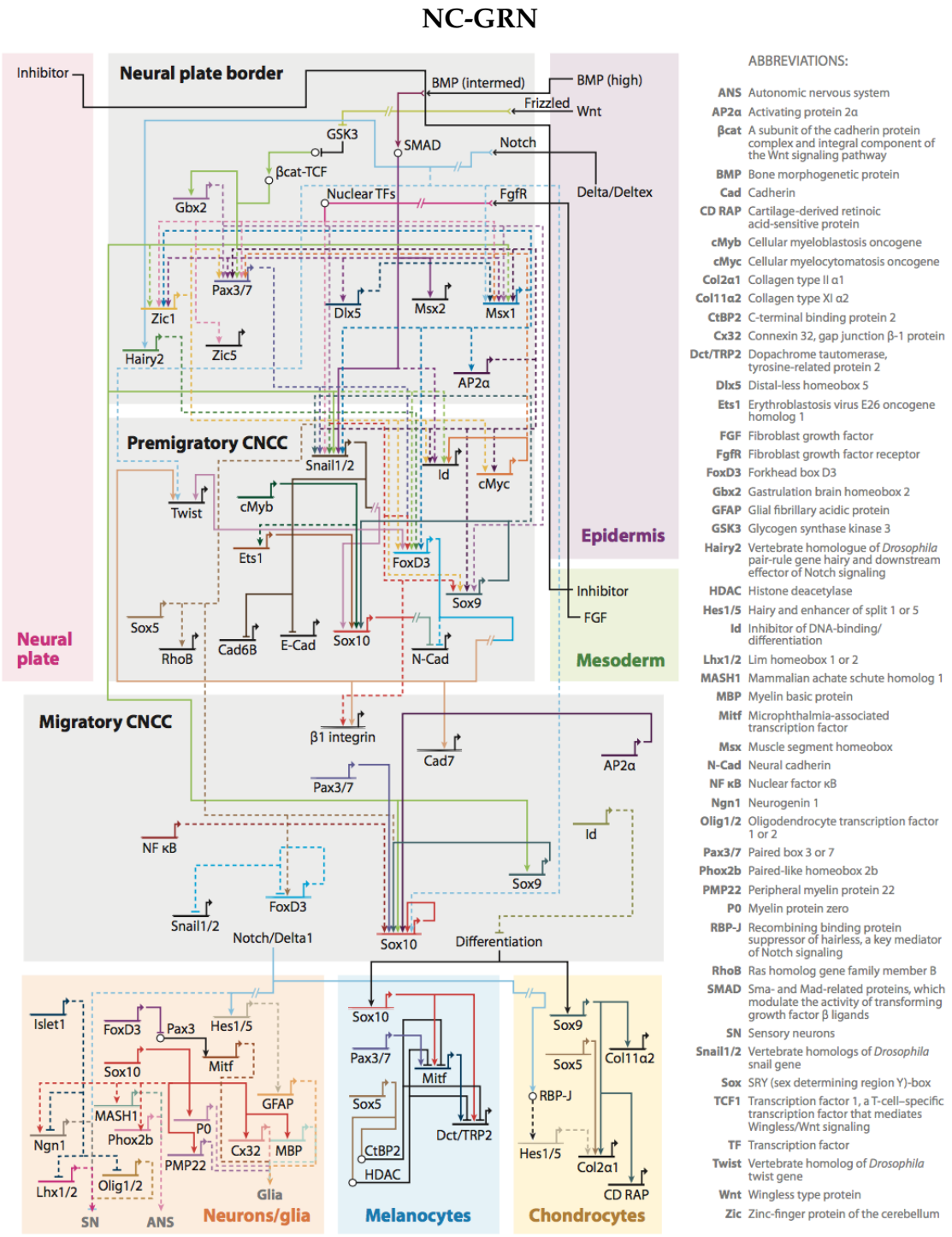


FIGURE 1.3: Diagram of the current NC-GRN detailing the known gene interactions at the neural plate border, pre-migratory, and migratory stages. Additionally, known network elements from derivative populations such as neurons/glia, melanocytes and chondrocytes are also shown here. Enhancer details are not shown in this diagram. Signalling inputs from the neural plate, epidermis and mesoderm at the neural plate border stage are also shown here. All abbreviations used within this diagram are detailed to the right of the network. (Image from (Betancur et al., 2010b))

For example, FoxD3 is one of the earliest markers of bona fide neural crest, is expressed by both pre-migratory and early migrating NC cells in all vertebrates studied to date (see Figure 1.4. Panels A & B) (Kos et al., 2001; Mundell and Labosky, 2011; Simões-Costa et al., 2012). During later stages of NC ontogeny, FoxD3 is down-regulated in a number of derivatives and has been shown to play a role in certain lineage decisions (Thomas and Erickson, 2009; Nitzan et al., 2013). In the chicken embryo, like in other vertebrates, exclusive expression of FoxD3 in the pre-migratory neural crest onsets in the cranial region. At this stage pre-migratory stage, tissue-specific RNA-sequencing shows FoxD3 transcripts account for 5-10% of the neural crest-enriched transcriptome (R.Williams unpublished data). FoxD3 is believed to be a pioneer factor, and as such does not require the pre-patterning of accessible and amenable chromatin states prior to engaging with its binding sites. FoxD3 binds across many sites globally and commences the EMT program within NCCs (R.Williams, unpublished data). FoxD3 expression in the cranial NC is faithfully recapitulated by the pre-EMT cranial NC specific enhancer, NC1 (see Figure 1.4.) (Simões-Costa et al., 2012). This NC- specific cranial enhancer is directly activated by Pax7, Msx1/2 and Ets1 transcription factors (see Figure 1.4. Panel C) (Simões-Costa et al., 2012). FoxD3 has an auto-regulatory negative feedback loop which utilises the NC1 enhancer. FoxD3 binds the NC1 element and leads to the shutdown of FoxD3 expression as the NC undergoes EMT. A mutated NC1 enhancer known as 'M3' can be used to maintain fluorescent reporter activity in migrating NC cells, as it does not contain the essential FoxD3 binding sites (see Figure 1.4. Panel D) (Simões-Costa et al., 2012).

1.3.2 Chromatin remodellers pattern global chromatin landscapes

During activate gene expression, chromatin must be opened such that TFs, co- activators and RNA polymerase II can access the DNA and initiate transcription. To ensure accessibility for these factors, the transcriptional start sites by active eukaryotic

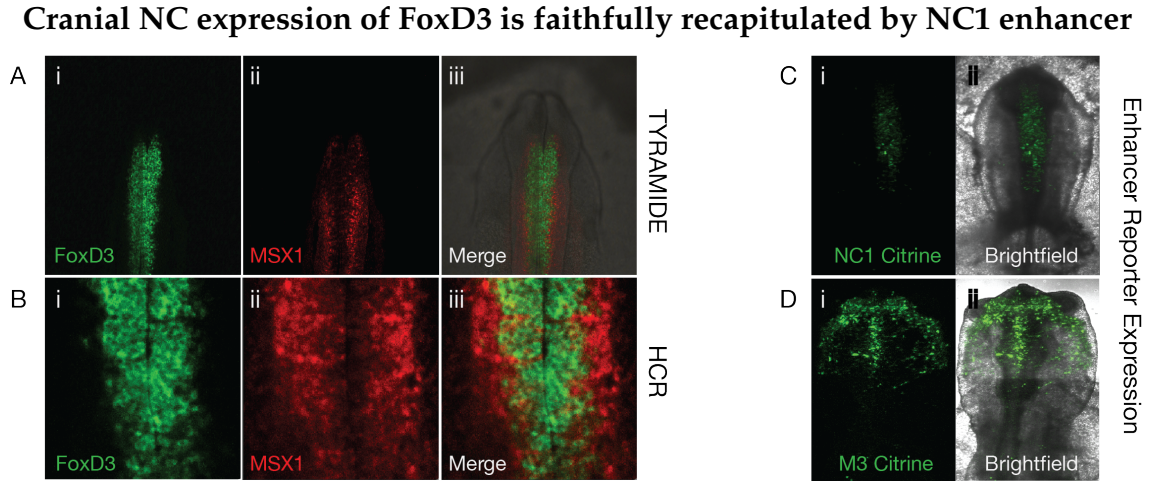


FIGURE 1.4: Confirmation of expression of FoxD3 is faithfully recapitulated by NC1 enhancer (Simões-Costa et al., 2012). (A) Tyramide *in situ* showing the expression pattern of bona fide pre-migratory NC TFs FoxD3 (i) and Msx1 (ii) in an early HH8 chicken embryo. (B) High-resolution image of HCR *in situ* for key NC TFs FoxD3 (i) and Msx1 (ii). (C) NC1 FoxD3 enhancer reporter activity in a late HH8 embryo highlighting the pre-migratory NC. (D) M3 FoxD3 enhancer reporter activity in a HH10 chicken embryo. The mutated M3 enhancer reporter activity is immune to the down-regulation experienced by NC1 and as such is clearly seen in migratory NC.

promoters are nucleosome-depleted. Conversely, stress response genes and developmentally regulated genes have increased nucleosome density at their promotors. This increased density interferes with accessibility of regulatory sequences, hampering transcription. Consequently, for transcription to commence or cease, chromatin remodellers, histone chaperones and other chromatin-modifying molecules must be employed to reorganise nucleosomes (Chen and Dent, 2014). TFs working in complex with chromatin remodelling and histone modifying factors are thought to ensure the site specificity of the remodelling and modifying activity.

The use of chromatin states to regulate gene expression is considered ubiquitous. As such, the components of chromatin machinery are ubiquitous also. Therefore, the disruption of genes coding essential and ubiquitous chromatin machinery should affect an entire organism. However, DNA methylation is ubiquitous eukaryotic mechanism which upon disruption results in remarkably restricted aberrations. For

example, CpG methylation is critical for X-inactivation, gene imprinting and embryonic development in general. These global DNA methylation patterns are established, and maintained, by the coordinated action of three ubiquitous DNA methyltransferases (DNMTs) (reviewed (Goll and Bestor, 2005)). DNMT3b is a *de novo* methylase involved in the establishment, but not maintenance, of these methylation marks. Despite ubiquitous low-level DNMT3b expression (Okano et al., 1999), hypomorphic DNMT3b mutations account for two-thirds of immunodeficiency, centromere instability, facial anomalies (ICF) syndrome cases(Maraschio et al., 1988). The facial aberrations in ICF syndrome, implicating NC hypersensitivity to DNMT3b loss-of-function, include hypertelorism, macroglossia, low- set ears, and epicanthal folds (Xu et al., 1999). Further evidence of NC sensitivity to DNMT3b activity is that ICF cells carrying DNMT3b mutations, in addition to the expected overall reduction of global methylation levels (Jeanpierre et al., 1993) additionally lack polycomb repression complex 2 (PRC2) signatures at homeobox genes critical for the development of the immune system, brain, and craniofacial region. Lentiviral knockdown of DNMT3b using a small hairpin RNA (shRNA) in H9 human embryonic stem cells alters the expression of early NC specifiers, and enhances premature NC precursor differentiation (Martins-Taylor et al., 2012). In addition to increased PAX7 and NGFR expression, DNMT3b knock-down cell lines also significantly over expressed PAX3, FOXD3, SOX10 and SNAI2, all highly significant to NC ontogeny, when compared to the mismatch shRNA expressing control cell line. This increased expression of regional identity specifying genes, and associated acceleration of differentiation, upon loss of DNMT3b is also seen in neural stem cells (Martins-Taylor et al., 2012). Taken together, the mutation of this ubiquitously expressed factor particularly reduces methylation at a limited range of sites. This suggests that DNMT3b, while a ubiquitous factor, may prefer certain genomic sites, which are of major significance only to certain tissues especially the NC. Along with CHD7, chromatin modifiers JMJD21 and DNMT2A are also reported to have important NC functions.

1.4 CHARGE syndrome: a Neurocristopathy

1.4.1 Neurocristopathies are characterised by aberrant NC development

Abnormalities in NC development are implicated in human syndromes and congenital conditions, collectively known as neurocristopathies (Bolande, 2010). These often feature assortments of congenital aberrations affecting multiple NC derived structures. Craniofacial anomalies are key features of DiGeorge, Waardenburg, Treacher Collins and CHARGE-syndromes respectively (Snider and Mishina, 2014), all of which are additionally associated with cleft-lip-palate phenotypes (Dixon et al., 2011). DiGeorge is a human primary immunodeficiency syndrome. It is characterised by T-cell deficiency/hypoparathoroidism, congenital heart disease, characteristic craniofacial tendencies and hypocalcemia. Waardenburg is a rare genetic condition characterised by hearing loss, minor defects in structures arising from the NC, changes to skin and eye pigmentation. Treacher Collins is a developmental defect characterised by aberrant development of bones and other tissues of the face. The contribution of the cranial NC to the musculo-skeletal foundation of the skull is of particular significance to all of these syndromes.

1.4.2 CHARGE syndrome

CHARGE syndrome (OMIM 214800) is a sporadic disorder, characterised by a multitude of congenital aberrations in tissues and structures (Schulz et al., 2014). 'CHARGE' is an acronym for the clinical features characterising the disease: ocular coloboma, heart malformations, atresia of the choanae, retardation of growth, genital hypoplasia and ear abnormalities. As these structures are significantly contributed to or patterned by the NC, CHARGE syndrome is considered a neurocristopathy. CHARGE syndrome affects 1:8,500-10,000 live births worldwide, with patient life expectancy

ranging from five days to 46 years. Approximately 30% of CHARGE syndrome patients die before the age of five (Zentner et al., 2010).

Clinical diagnosis is based on major and minor diagnostic criteria. There is controversy surrounding which are the major clinical characteristics and how many criteria must be met to result in a clinical diagnosis (Blake et al., 1998; Verloes, 2005). The phenotypic aberrations in CHARGE syndrome patients can vary in both severity and presentation (Basson and van Ravenswaaij-Arts, 2015). For example, while 76% of patients have congenital heart defects, choanal atresia only occurs in 55% of patients (Bergman et al., 2011). Interestingly, there is a striking level of phenotypic variability, even in rare familial case where the same mutation is carried.

CHARGE syndrome is particularly characterised by aberrations in a number of structures where key interactions between the NC and other cell types are known to occur (see Figure 1.5.). CHARGE syndrome is increasingly thought of as a neurocristopathy as evidence gathers demonstrating the aberrant migration of NC in model systems recapitulates the aetiology of CHARGE syndrome (Siebert et al., 1985). Interestingly, during development, primordial germ cells (PGCs) are thought to migrate into the gonads from the mesentery of the hindgut, specifically by interacting with NC- derived Schwann cells (Møllgård et al., 2010; Le Douarin and Teillet, 1973). PGCs go on to interact extensively with steroidogenic support cells within the developing gonad (Jemc, 2011). If PGC migration is halted or impaired, the steroidogenic supporting cells of the gonad may not secrete appropriate hormones. Thus, while the NC contribution to the genital hypoplasia is evidently not direct, CHD7-attributable aberrant NCC migration or differentiation could well contribute to the aetiology of CHARGE syndrome even in this context.

The 'minor' features of CHARGE syndrome include hypoplasia of the semicircular canals, cranial nerve abnormalities and facial nerve palsy, impaired cognitive function, tracheoesophageal stula and fistulas, and distinct facial asymmetries including cleft lip/palate (Bergman et al., 2011; Jongmans et al., 2006; Lalani et al., 2006; Sanlaville and Verloes, 2007; Verloes, 2005). 80-90% of CHARGE patients have

anosmia and cleft lip/palate related feeding problems. Additionally, malformation of the foregut, reproductive organs, kidneys, limbs and abnormal spinal curvatures, digits, brain, pituitary gland, lung abscesses, focal hepatic necrosis and immune system problems have also been observed in CHARGE patients (Blake et al., 1998; Sanlaville and Verloes, 2007).

1.4.3 Mutations in chromatin remodeller CHD7 cause CHARGE syndrome

CHARGE patients present with aberrations of structures contributed to, or patterned by, NCCs. 60% of CHARGE syndrome cases are accounted for by mutations in CHD7 (Vissers et al., 2004). Studies have reported over 500 different pathogenic mutations throughout the CHD7 gene body, suggesting premature termination at any point of the protein is significantly detrimental to CHD7 function (Basson and van Ravenswaaij-Arts, 2015; Bergman et al., 2011; Bilan et al., 2012; Janssen et al., 2012; Vissers et al., 2004; Vuorela et al., 2007). The remaining 40% of cases, which do not have mutations within the CHD7 gene body, could be attributable to mutations in elements or regulators downstream of CHD7, or in upstream *cis*-regulatory regions of CHD7 (Bergman et al., 2011; Zentner et al., 2010).

The CHD family of chromatin remodellers contains 12 highly conserved members (Woodage et al., 1997; Hargreaves and Crabtree, 2011).The CHD family is defined by two motifs: tandem N-terminal chromodomains, responsible for binding methylated histones (Brehm et al., 2004) and a central SNF2-like ATPase/helicase domain containing chromatin-remodelling activity. The major sequence variation between family members is found in the C-terminal region known to contain the DNA binding activity (Zentner et al., 2010).

As a member of the chromodomain helicase DNA-binding family of ATP-dependent chromatin remodelling enzymes, CHD7 catalyses the repositioning of nucleosomes along chromatinised DNA (Chen and Dent, 2014). CHD7 chromatin remodelling

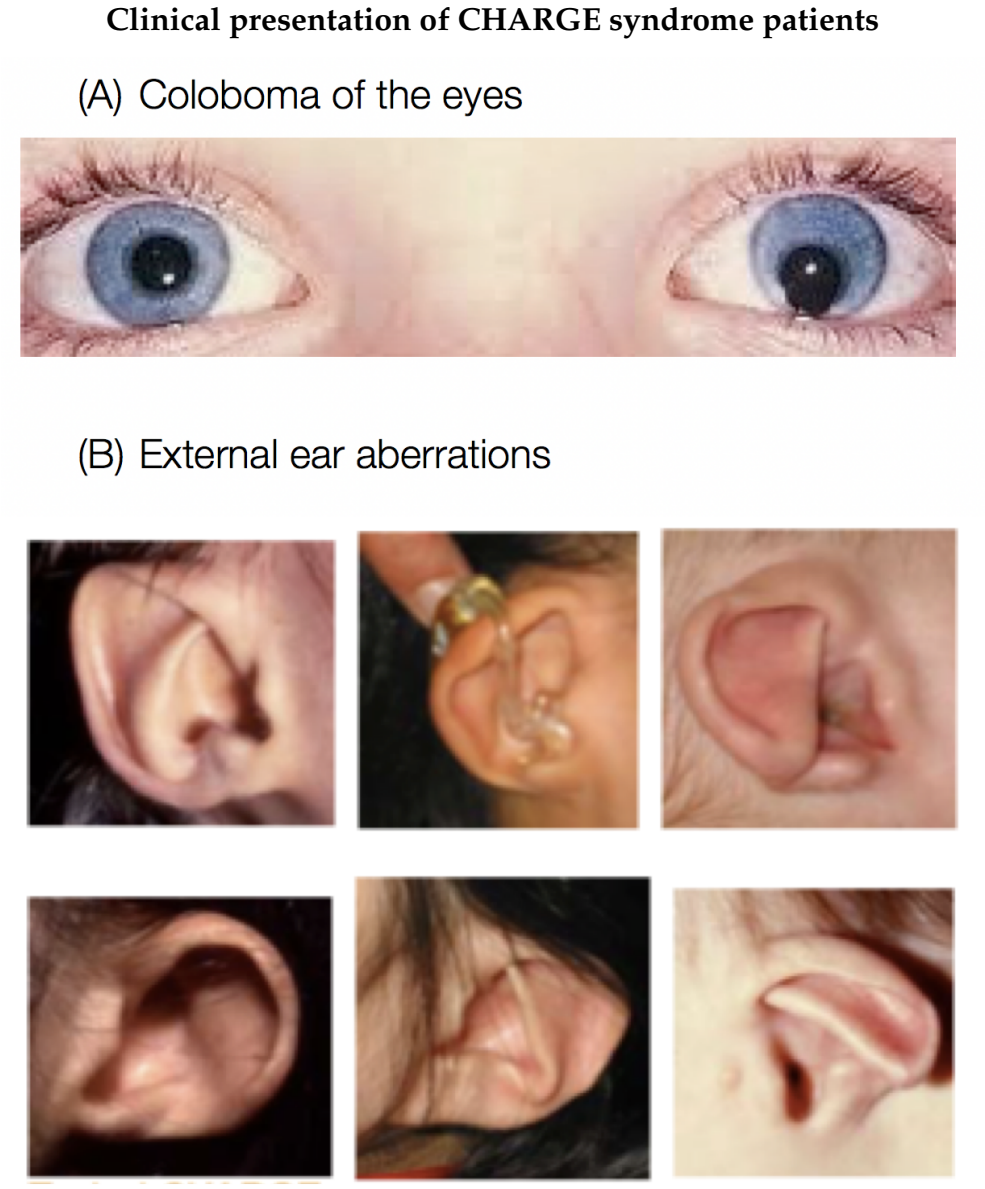


FIGURE 1.5: Images showing characteristic CHARGE syndrome cranial structures. (A) coloboma of the eye (unilateral, left eye). (B) external ear defects. Internal ear defects not shown here.

activity is important for normal gene regulation, as specific nucleosome positioning at regulatory regions is critical for productive associations of TFs, core transcription machinery and DNA (Basson and van Ravenswaaij-Arts, 2015; Clapier and Cairns, 2009). The CHD7 double chromodomain is used to couple ATP-hydrolysis derived energy into nucleosome repositioning, and mediates CHD7 recruitment to enhancers by binding specific histone methylation marks (Narlikar et al., 2013) ChIP-sequencing in ES cells has shown that CHD7 is preferentially recruited to sites marked by H3K4me1 histone modifications. H3K4me1 marks poised enhancers, and together with H3K27ac, is also found on active enhancers (Visel, Axel et al., 2009; Creyghton et al., 2010). Additionally, this binding distribution dynamically reflects changing H3K4me1 and H3K4me2 profiles, suggesting its recruitment is related to the methylation state at enhancers (Schnetz et al., 2010, 2009). While cell type-specific enhancers may underlie the context-specific role of CHD7 (Basson and van Ravenswaaij-Arts, 2015), CHARGE syndrome has been suggested to arise due to CHD7 regulated enhancer-mediated gene dysregulation (Schnetz et al., 2010).

CHD7 contains two chromodomains which are used to bind histone methylation marks and help couple ATP hydrolysis derived energy into nucleosome repositioning (Narlikar et al., 2013). The HELICc domain hydrolyses ATP and drives RNA duplex unwinding. This is thought to generate the force to translocate itself along DNA, moving protein complexes such as nucleosomes out of the way (Bork and Koonin, 1998; Sengoku et al., 2006; Langst and Manelyte, 2015). The SANT domain is thought to be involved in the protein-protein interaction of CHD7 and the histones. The BRK domain translocates nucleosomes along DNA. Finally, CHD7 contains five nuclear localisation sequences, functionally confirmed to localise CHD7 to the nucleus (see Figure 1.6.) (Schnetz et al., 2009)

In mouse ESCs CHD7 is found at sites with high H3K4me1 histone modifications, a key characteristic of enhancers, together with Brg1 (Schnetz et al., 2009; Bajpai et al., 2010). Proteins involved in chromatin remodelling are typically found in multi-protein complexes. CHD7 co-localises with p300, Oct4, Sox2, Nanog, Smad1

and Stat3 at enhancer elements in murine ESCs (Schnetz et al., 2010). This lead to the hypothesis that these proteins are co-factors in enhancer-promoter interactions, and thus possible that CHARGE syndrome may arise due to CHD7 enhancer-mediated gene dysregulation. CHD7 is expressed in ESCs (Schnetz et al., 2009). Sox2 is also expressed in ESCs and is one of the four key TFs involved in iPS cell reprogramming. CHD7 significantly interacts with Sox2 in ESCs (van den Berg et al., 2010). Furthermore, mutations of SOX2 result in anophthalmia syndrome (Fantes et al., 2003), the phenotypic features of which significantly overlap with that of CHARGE syndrome. Misdiagnosis of SOX2 patients as CHARGE patients has been noted within the literature (Williamson et al., 2006). Direct Sox2 and CHD7 target sites overlap significantly in ESCs (Schnetz et al., 2009), and neural stem cells (NSCs) (Engelen et al., 2011). Interestingly, mutations affecting a subset of these CHD7/ SOX2 bound targets in NSCs cause human syndromes with phenotypic similarities to CHARGE and anophthalmia syndromes respectively. For example, Jagged 1 (Jag1) is bound by SOX2 and CHD7 in NSCs, Jag1 mutations cause Alagille syndrome which is characterised by semi-circular canal defects- which is also 100% penetrant in CHARGE syndrome. A reduction in Jag1 results in this phenotype be it from Jag1 mutations or a reduced CHD7 expression level impacting that of Jag1 (Engelen et al., 2011). This suggests that the gene regulatory networks regulated by SOX2 and CHD7 factors overlap at points, and that the similarities in the human syndrome traits reflect this hierarchy.

Mutations within the exons of CHD7, situated on human chromosome 8, lead to haploinsufficiency and account for up to 70% of CHARGE cases studied. Most reported mutations lead to truncation of the mature protein. Nonsense mutations (44%) and frame shift indels (34%) account for the most frequent type of pathogenic variants in CHARGE patients. There are reports that exons 2, 8, 31 and 34 were most affected by mutations in patients with CHARGE. However, over 500 different pathogenic CHD7 mutations have been reported across the entire gene, not only in the functional protein domains (see Figure 1.6.). Homozygous CHD7 mutations are

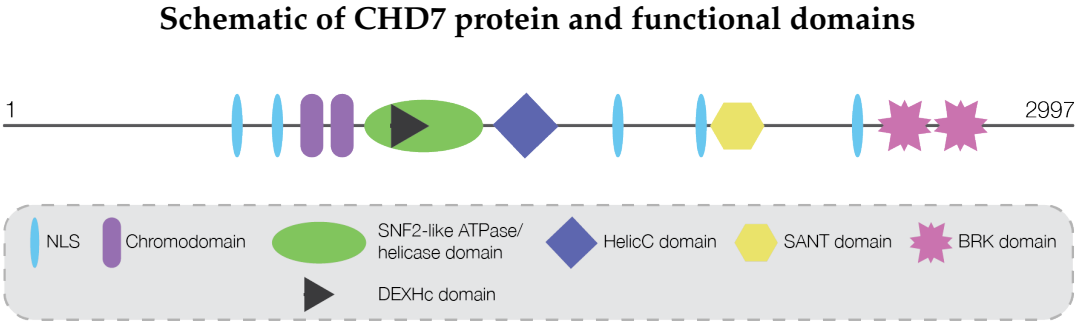


FIGURE 1.6: CHD7 is located on chromosome 8q12.1, has 38 exons and spans 188kb. CHD7 has tandem N-terminal chromodomains (purple), a central helicase domain (green) characteristic of CHD family proteins, and a HELICc domain (navy) to bind ATP. It also contains a C-terminal DNA binding domain that shares a low degree of homology with the histone/DNA-binding SANT domain (yellow), and two C-terminal BRK domains (pink) with unknown functions. Nuclear localisation is ensured by five consensus nuclear localisation signals (NLS) (blue). Figure adapted from (Zentner et al., 2010).

embryonically lethal with CHD7 null gene-trap mice dying at E10.5 (Hurd et al., 2007).This lethality is attributed to defects within cardiovascular structures. This is additionally supported by the finding that CHD7 expression onsets at E7 in the vascular plexus (Bosman et al., 2005). As for other Chd7 null alleles, homozygous $Chd7^{Gt/Gt}$ mice do not survive beyond E10.5, likely due to abnormalities in cardiovascular structures (Bosman et al., 2005; Hurd et al., 2007).

1.5 Fitting Chromatin Remodeller CHD7 into the Cranial NC-GRN

In human NC-like cells CHD7 has been shown to be associated with components of the BAF- (Brahma associated factor complex) and PBAF - complexes (Polybromo containing complex) (Bajpai et al., 2010). Both belong to the SWI/SNF-family of ATP-dependent chromatin remodelling complexes and can act as transcriptional activators or repressors (Chen and Dent, 2014; Wang et al., 2007). In both humans and Xenopus, CHD7 is essential for NC migration but not NC induction (Bajpai et al., 2010). In Xenopus, CHD7 is upstream of key NC TFs Sox9, Twist and Slug/Snail2,

as demonstrated by a reduction in expression levels upon CHD7 knockdown and respective rescue by mRNA and MO co-injection experiments (Bajpai et al., 2010).

Knockdown of CHD7 in Xenopus embryos recapitulates all major features of CHARGE syndrome in humans (Bajpai et al., 2010) and mice with heterozygous loss of function mutations in Chd7 are a model of CHARGE syndrome (Layman et al., 2010). CHD7 is required for activation of critical NC transcriptional circuitry, including Sox9, Twist, and Slug (Bajpai et al., 2010). CHD7 is responsible for chromatin remodelling and is widely expressed during development, helping to explain the broad range of congenital defects observed. The role of CHD7 in chromatin remodelling reinforces the emerging view that epigenetic regulation plays an important role in NCC development (Liu and Xiao, 2011; Keyte and Hutson, 2012). Additionally, CHD7 regulates osteoblastogenesis process involved in the formation of the craniofacial region by NC- derivatives (reviewed by Dupin and Sommer (2012)).

In conclusion, the process and outputs of NC development are highly complex, therefore requiring complex spatio-temporal regulation to ensure this level of complexity is both achieved and faithfully maintained. This observable complexity must be driven by an underlying regulatory network with precise spatio-temporal regulation mediated by enhancers. Additionally, at the point of EMT, NCCs undergo large-scale chromatin landscape priming to allow their significant change in morphology and behaviour. It is conceivable that CHD7 might be the chromatin remodeller of choice for the NC. The clinical presentation of CHARGE syndrome patients further indicates that the NC is particularly reliant on the function of this chromatin remodeller. While CHD7 expression is clearly important to the NC as a whole, the cranial sub-population appears most sensitive to perturbations. Unpublished work in the Sauka-Spengler lab has shown that cranial NC cells express high levels of CHD7 compared to non-NC cells in the cranial region. This further supports the hypothesis that CHD7 has a significant role in the cranial NC. To place CHD7 within the cranial NC-GRN requires the identification of NC specific enhancers which are bound by NC TFs and ensure expression of CHD7 specifically in NCCs. Much is

already known about the cranial NC-GRN (see Figure 1.3.), consequently, if any key NC TFs bind CHD7 enhancers, it will be possible to fit CHD7 into the established GRN.

1.5.1 Identifying enhancers using the classic chicken embryo model system

The chicken is a classical developmental biology system as it is amenable to grafting experiments, lineage tracing and can be manipulated at a number of stages both *in-* and *ex-ovo* (Stern, 2005). Whilst the chicken system has been widely used to address gene expression patterns and functional comparisons, it is not particularly amenable to genetic manipulations (Meulemans and Bronner-Fraser, 2004). Consequently, experimental probing of the NC-GRN has been chiefly restricted to loss- and gain-of-function involving cDNA, mRNA, morpholino or antisense oligonucleotide injections (Sauka-Spengler and Barembaum, 2008). Whilst the development of highly efficient electroporation techniques has addressed previous issues surrounding mosaic reporter expression assays, the difficulty of generating genome targeted modifications in an efficient fashion have been lacking.

For example, the pre-migratory cranial NC FoxD3 enhancer was identified by conservation analysis. Eighteen conserved regions were cloned into EGFP reporter vectors and electroporated into HH4 chicken embryos. These reporter constructs contained minimal promoters upstream of the EGFP coding region but require enhancing to drive reporter signal. A ubiquitous tracer, pCI-H2B-RFP was used to confirm transfection efficiency. Once a region was identified as having activity, in the case of NC1 888bp, progressive cloning of smaller and smaller fragments was tested to see if they recapitulated the expression initially observed. This allowed the isolation of a 553bp minimal NC1 enhancer region (Simões-Costa et al., 2012).

1.5.2 Transcription Factor binding site profiling in chicken

Addressing questions of direct TF binding in the NC has been a technical challenge in the chicken has the number of cells is very limited, as is the affinity and specificity of TFs. However, binding information is required for the resolution of the NC-GRN, given overexpression and knockdown effects are insufficient to establish direct interactions.

Our lab has developed a genetically-encoded binary approach called 'biotagging', which utilises *in vivo* biotinylation as previously described in other model organisms (de Boer et al., 2003; Deal and Henikoff, 2011) to enable regulatory profiling and characterisation of molecular interactions in zebrafish embryos. By co-expressing a biotin acceptor peptide (Avi-) tagged TF-of-interest and a bacterial biotin ligase (BirA), the addition of a single biotin residue by BirA to the Avi-tagged TF-of-interest allows the use of downstream streptavidin-based affinity purification methods to isolate and characterise DNA elements bound by the TF-of-interest (Kim et al., 2009; Schatz, 1993; Beckett et al., 1999; de Boer et al., 2003). This method allows high stringency washing, reducing background signal, and also permits efficient purification of TF-DNA complexes from small cell populations. As such, if translated into the chicken system, it could be a valuable tool in the resolution of the NC-GRN by allowing dissection of direct TF binding *in vivo*.

1.5.3 Genome- and epi-genome editing technologies in chicken

The advent of CRISPR-Cas9 mediated genome engineering has provided the tools to begin addressing this issue of transgenesis within the chicken system. This system relies on a synthetic guide RNA (gRNA) to direct the activity of the Cas9 endonuclease to user-defined target sites (see Figure 1.7.). Upon targeting, the Cas9 endonuclease causes a double strand break, allowing the possibility of a DNA repair mechanism introduced error to disrupt the targeted sequence (Jinek et al., 2012;

Cong et al., 2013). This genome editing technology provides the means to do improved experiments and to verify interactions between factors, thus improving our ability to dissect and resolve GRNs during embryonic development. Additionally, the rate of advanced CRISPR- Cas9 modifications, such as activator and repressor module tethering to dead-Cas9 (dCas9), in other fields is rapidly advancing. As such, there is a vacancy for the adaption of CRISPR-Cas9 techniques for use in the chicken system that needs to be addressed.

In chicken, the first proof-of-principle *in vivo* CRISPR-mediated somatic-cell genome engineering experiments were reported in 2015. Veron *et al* presented effective Pax7 loss-of-function in the somite of chicken embryos, as measured by immunofluorescence (Veron et al., 2015). Subsequently, there have been a handful of additional papers applying genome editing and further developing CRISPR-Cas9 mediated tools for use in the chicken system (Kwaku Dad Abu-Bonsrah and Newgreen, 2016; Shashank Gandhi and Bronner, 2016), including work presented here (Williams et al., 2018).

1.6 Project Summary & Objectives

NC derivatives contribute a wide range of cell-types and tissues to the vertebrate embryo. Consequently, aberrant NC development impacts a wide range of structures containing NCC derivatives. Neurocristopathies are pathologies characterised by malformations of structures containing significant NC contributions. CHARGE syndrome is a neurocristopathy where approximately 60% of patients are heterozygous for deleterious mutations affecting the exons of chromatin remodeller CHD7. It is surprising that the loss of a ubiquitously expressed chromatin regulator like CHD7 has the limited, and ostensibly tissue-specific, effect seen in human CHARGE-syndrome. However, increasing evidence from model organisms and human iPSC derived NCLCs demonstrates NCC migration is significantly affected by aberrations in CHD7 expression. CHD7, together with other chromatin regulators KDM4A

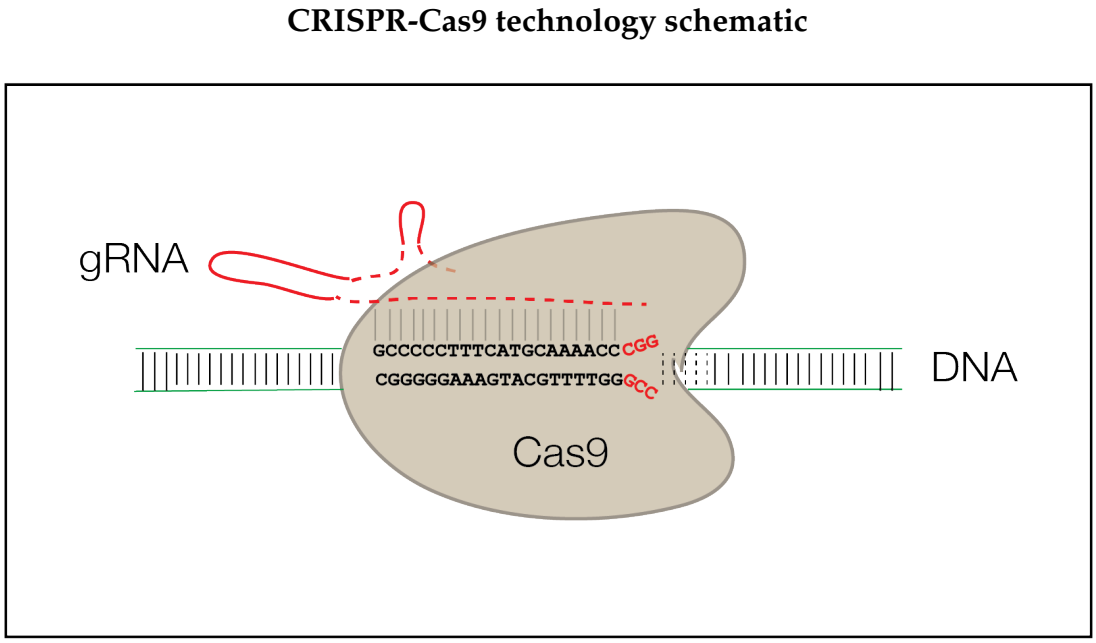


FIGURE 1.7: The gRNA (red) is loaded into the Cas9 endonuclease protein (beige). The 'spacer' sequence targets Cas9 to a specific DNA sequence (black). The target DNA sequence is edited as a result of the Cas9 endonuclease activity and the cell's DNA repair mechanisms.

and DNMT3A, has been shown to have high levels of expression at early stages of NC ontogeny prior to EMT (Hu et al., 2012; Bajpai et al., 2010). During the differentiation of hESCs into NC progenitors or neurectoderm progenitors, the expression of these factors increases over fivefold. Of note, the two different progenitor populations showed clear and differential preference for which chromatin regulator they upregulated (Weng et al., 2012). This spatio- temporally specific and differential expression implicates the presence of enhancers, which may be utilised by different tissue-specific factors to regulate CHD7 expression.

The prospect that CHD7 may have tissue-specific enhancer elements both potentially explains CHARGE patients who do not have mutations within the exons of CHD7, but also would allow the integration of a chromatin remodeller into the NC-GRN. To rationalise the activity of CHD7 within the NC and fully integrate this chromatin remodeller into the NC-GRN, requires identifying the enhancers and

the upstream factors regulating CHD7 expression via these elements. Given different axial levels of the NC are known to be regulated differently, and CHARGE syndrome aetiology implicates the cranial NC in particular, we sought to integrate CHD7 specifically into the cranial NC-GRN.

As such, I intend to investigate the possible NC-specific regulation of chromatin remodeller Chd7 during chicken embryonic development. This will require the development of an advanced toolkit for the exploration of gene regulatory interactions in the small cell population of the NC. Thus, I will adapt the Sauka-Spengler zebrafish 'biotagging' technique for use in profiling TF binding, and additionally develop the translation of CRISPR-Cas9 technologies into chicken embryos. Taken together, this will allow the in-depth characterisation of any identified novel enhancers, and functional interrogation of the direct links established using these techniques, and improved GRN resolution using the chicken model system.

Given our intention to locate and characterise cranial NC-specific enhancers for Chd7 using the chicken model system the following objectives must be met:

Objective 1: Establish an advanced molecular toolkit to probe gene regulatory networks by translating biotagging of transcription factors and CRISPR-Cas9 based genome- and epigenome- engineering techniques for use in the chicken model system.

Objective 2: Examine the genomic locus surrounding chromatin remodeller Chd7 for tissue-specific regulatory elements and screen putative elements for enhancer activity *in vivo*.

Objective 3: Employ the advanced molecular toolkit to characterise novel Chd7 tissue- specific enhancers and place within the cranial NC-GRN.

Chapter 2

Materials & Methods

2.1 Chicken Embryo Methods

Fertilised wild type chicken eggs were obtained from Henry Stewart & Co (Norfolk). All chicken embryos were handled in accordance with UK Home Office regulations and local ethical approval from the University of Oxford.

2.1.1 Chicken embryo incubation

Fertilised chicken eggs were incubated at 37°C in a humidified incubator (Brinsea Ova- Easy Advance incubator).

2.1.2 Chicken embryo collection and electroporation

Electroporation was used to introduce experimental plasmids into the embryonic epiblast for downstream investigation. Prior to electroporation, embryos were incubated for circa 22 hours until Hamburger & Hamilton stage 4 (Hamburger and Hamilton, 1951) was reached. Embryos were manually harvested on Whatman filter paper rings, cleaned in Ringer's solution, and electroporated as previously reported (Sauka-Spengler and Barembaum, 2008). Embryos were placed ventral-side up in an electroporation chamber, filled with Ringer's solution, above a platinum plate electrode at the bottom of the central well. All constructs were injected into the lumen between the vitelline membrane and the epiblast using a glass needle.

A square pulse electroporator was used to apply five 5-Volt 50 millisecond pulses, separated by 100 millisecond pauses, via a flat-tip platinum electrode to the entire epiblast. Embryos were grown on 1mL fresh thin albumin containing penstrep antibiotic (1:1000 concentration) in a small petri dish at 37°C overnight until the desired Hamburger & Hamilton stage was reached to observe fluorescence (Betancur et al., 2010b; Hamburger and Hamilton, 1951; Sauka-Spengler and Barembaum, 2008).

2.2 Putative *Cis*-Regulatory Element Selection and Enhancer Testing Methods

2.2.1 Identification of putative *cis*-regulatory elements

To identify putative NC-specific *cis*-regulatory elements for Chd7, ATAC-sequencing profiles from FACS purified NCCs from embryos at 4ss and 9-12ss were compared to FACS negative cells. FACS purification was conducted using an NC1-FoxD3 enhancer driven fluorophore, thus selecting for bone fide specified NCCs. NC1 is a published FoxD3 enhancer, with activity reported from pre-EMT into early EMT stages (Simões-Costa et al., 2012). ATAC-sequencing is a transposase-dependent method of profiling chromatin accessibility in cell populations (Buenrostro et al., 2013). Next- generation sequencing adaptors are introduced by a transposase into accessible sites of the genome. This results in an increased density of sequencing reads in regions of open chromatin compared to closed chromatin. Regulatory elements, such as active enhancers, are accessible to regulator binding and are consequently sensitive to transposase activity. Thus, active promotors and enhancers appear as peaks in ATAC-sequencing profiles.

The genomic locus encompassing Chd7, from the end of the neighbouring upstream gene (RAB2A chr2: 112,558,232) to the start of the downstream gene (Clvs1 chr2: 112,882,670), was surveyed for peaks. Peaks which had dynamic increases in

height from the 4ss to 9-12ss stages were deemed putative *cis*-regulatory sites potentially harbouring Chd7 NC-specific enhancers.

2.2.2 Functional testing of putative enhancers

To confirm putative Chd7 enhancer activity, the DNA sequence was cloned upstream of a minimal promoter and a fluorescent reporter. This minimal promoter is incapable of driving fluorophore expression unaided, thus the inclusion of additional DNA tests for the enhancing capacity of the cloned sequence. The plasmid containing the putative enhancer can then be electroporated into the chicken embryo epiblast and monitored for fluorescence over time. The assumption is, if fluorescence is observed, the factors which bind the cloned sequence are present in that population of cells, and together they are able to enhance the activity of a downstream gene.

Therefore, to demonstrate enhancing potential of putative Chd7 enhancers, each individual region was cloned into a vector containing a minimal promoter upstream of a fluorescent reporter. If expression was observed, it was deemed a positive enhancer, and the potential population utilizing the novel enhancer was inferred from the spatio- temporal pattern of reporter activity. This pipeline was used for all putative and mutation site screening enhancers cloned and tested (see Appendix 1 for details).

Fluorescent reporter vectors containing Citrine, Cerulean or mCherry were used to test individual putative enhancer sequences. GoldenGate molecular cloning was used to simultaneously and directionally introduce the putative enhancer fragments and ligate them into the vector backbone. GoldenGate cloning relies on the activity of Type II restriction enzymes such as BsmBI to create these specific overhangs, and T4 DNA ligase to reconstitute the vector. Each vector has specific non-palindromic overhangs, subsequent to cleavage by Type II enzyme BsmBI, which differentiate them from each other. This allows the specific integration of PCR products with

complementary tails to these overhangs by T4 DNA ligase in the GoldenGate reaction. GoldenGate compatible primers were designed to span putative enhancer regions. These PCR products were generated using long-range DNA polymerases (KAPA Biosystems, cat #KK2601, or NEB, cat #M0532), then gel purified (Promega, cat #A9282) prior to the GoldenGate BsmBI reaction. 75ng of PCR product was added to 75ng of reporter vector plasmid, 1μL BsmBI (NEB cat #R0580), 2μL 10x T4 Ligase Buffer, 1μL T4 Ligase (NEB cat# M0202S) in a 20μL mixture of H_2O. The GoldenGate reaction was as follows: 25 cycles of 37°C 5 minutes, 16°C for 5 minutes, 55°C 5 minutes, 80°C 5 minutes. Contaminating bacterial chromosomal DNA was removed using Plasmid- SafeTM (Epicentre, cat #E3101K) prior to transformation into competent cells. IPTG-/X- gal selection on ampicillin plates was used to initially screen for positive integration of PCR products into reporter vectors. Colonies were size screened on agarose gels, miniprepped and sequenced to identify putative enhancer positive clones.

2.3 Whole-Mount Imaging Methods

2.3.1 Labelled RNA probe generation

DIG- or FITC- labelled anti-sense probes were generated against mRNAs of interest for downstream antibody detection and subsequent visualization of mRNA localization in an embryo. The templates for FoxD3, Sox10 and Chd7 anti-sense probes were PCR generated using cDNA from 6-12 somite chicken embryos (see Appendix 5 for details). Subsequent to confirming the size of the PCR products, the fragments were gel- extracted and further purified using phenol-chloroform extraction. T7 RNA polymerase (Promega, #P207B) was used to transcribe 1μg of template PCR product at 37°C in 2- 3 hours. Digoxygenin (DIG)-labelled dNTPs (Roche, #12430721) or fluorescein (FITC)- labelled dNTPs (Roche,

#11685619910) were used to label the probes. Probes were doubly purified using G-50 micro-columns (GE Healthcare, #28-9034-08).

2.3.2 Wholemount *in situ* hybridisation

To highlight the location of an mRNA of interest in an embryo, DIG-labelled probes were detected by enzyme-conjugated antibodies and used to generate coloured precipitates at the site of mRNA expression. Embryos were fixed in 4% PFA in phosphate-buffered saline (PBS) (0.1 M, pH 7.4) for an hour at room temperature (RT) and washed in DEPC-PBT (DEPC treated 1x PBS + 0.1% Tween-20), prior to dehydration using a methanol series and storage at -20 °C in 100% methanol till use. Embryos for use were serially rehydrated to DEPC-PBT and then equilibrated with hybridization buffer (50% Formamide, 1.3 mM SSC pH 5.0, 5 mM EDTA pH 8.0, 200 μg/mL Baker's yeast tRNA, 0.2% Tween-20, 0.5% 3-[(3- cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), 100 μg/mL Heparin). Embryos were incubated overnight with DIG-labelled probes in hybridization buffer at 65°C. Subsequent to RNA-probe removal and storage for reuse, embryos were washed in MABT (4 x 30mins) (1x MAB +0.1% Tween-20). Embryos were then blocked with 20% sheep serum +2% Boehringer Blocking Reagent (Roche, cat # 11096176001) in MABT. Following overnight incubation at 4°C with anti-DIG-alkaline phosphatase antibody (Roche, cat # 11093274910) excess secondary antibody was washed using MABT (at RT: 2x 30mins, 6x 1hr, at 4°C: 1x overnight). Signal was developed using BM Purple AP (Roche, cat # 11442074001). Once desired colour was reached, 2 x100mM Tris pH 9.5 washes and 3x PBT washes were used to stop the reaction. Light microscopy was used to whole-mount image the embryos as previously done within the lab (Betancur et al., 2011). If images could not be immediately taken upon the Olympus MVX upright microscope, post-fixation and dehydration was conducted again prior to re-imaging.

2.3.3 Tyramide fluorescent *in situ* hybridization

Tyramide fluorescent *in situ* hybridization was used to fluorescently report the location of multiple mRNAs in an embryo, prior to imaging using a fluorescent microscope. Embryos were collected and prepared as in the wholemount *in situ* hybridization protocol (see above). Embryos were rehydrated from 100% Methanol to 1x PBT (1x PBS + 0.1% Tween-20) using a graduated series of these solutions. Endogenous peroxidases were inactivated by incubating embryos in 1% H_2O_2 for 15 minutes and then washed in PBT twice. Embryos were then incubated in hybridization buffer 50% (as in wholemount *in situ* hybridisation) for an hour prior to probe addition. Embryos were incubated at 65oC overnight in hybridization buffer 50% and probes. Subsequent to probe recover, embryos were washed in KTBT (50mM Tris-HCl pH7.5, 150mM NaCl, 10mM KCl and 0.2% Tween-20) at RT (4x 30 mins), and blocked in 15% sheep serum + 0.7% BBR in KTBT for 1 hour at RT. DIG-labelled probes were detected using anti-DIG-POD (Roche, cat #11207733910) at 1:1000 dilution in blocking solution overnight at 4°C. After washing excess antibody away using KTBT (3x 5 minutes, 10x 20 minutes), signal was amplified using Tyramide-Cy3 reagent (Perkin Elmer cat # NEL53001KT). To inactivate the remaining anti-DIG-POD activity, samples were incubated in 1% H_2O_2 in KTBT for 45 minutes at RT. FITC-labelled probes were then detected using anti-FITC-HRP (Roche, cat #11426346910) at 1:1000 dilution in blocking solution overnight at 4°C. After washing excess antibody away using KTBT (3x 5 minutes, 10x 20 minutes), signal was amplified using Tyramide-FITC reagent (PerkinElmer cat # NEL53001KT). Prior to imaging, HRP inactivation was conducted by incubation in 1% H_2O_2 in KTBT for 45 minutes at RT and 2x PBS washes. Images were taken on both inverted and upright Zeiss 780 confocal microscopes (10x or 25x objectives).

2.3.4 Hybridisation Chain Reaction (HCR) fluorescent *in situ*

Hybridization HCR is a fluorescent *in situ* hybridisation technique which utilises programmable amplifiers to allow highly multiplexed target mRNA detection. Short probes are designed to tile across the target mRNAs, and each carry a tail which specifically binds a fluorescently-labelled hairpin. Upon binding to the probe, this hairpin loses the secondary structure and becomes a similar binding site for a second fluorescently- labelled hairpin. This triggers a chain reaction, where two fluorescently-labelled hairpin species alternatingly bind each other and form long, mRNA-tethered chains (Dirks and Pierce, 2004; Choi et al., 2016). There are a number of different species of amplifier, sets of two hairpins carrying the same fluorophore, which allow multiplexed mRNA detection in samples.

2.3.5 Whole-mount immunocytochemistry

Antibodies conjugated to fluorescent molecules were employed to bind and report on the localization of proteins of interest, as previously done within the lab (Betancur et al., 2011). Embryos were fixed in 4% PFA in phosphate-buffered saline (PBS) (0.1 M, pH 7.4) for 1 hour at RT. Embryos were then rinsed for 3x 15 minutes to 1 hour in PBT (2% DMSO, 0.5% Triton in PBS), incubated in block (10% donkey serum in PBT) for 2 hours at RT. Overnight incubation with primary antibody (1:200 in block) was at 4°C. Excess primary antibody was washed with PBT (3-5x 1hour at RT, at 4°C overnight). Embryos were incubated in secondary antibody (1:500 in PBT) for 2 hours at RT and underwent a final series of PBT washes (6-8x 1 hour at RT and 4°C overnight) prior to imaging. Images were taken on both inverted and upright Zeiss 780 confocal microscopes (10x or 25x objectives). DAPI in mounting PBS solution was at 1:20,000. Primary antibodies: 1:500 rabbit anti-GFP-488 direct conjugate (Abcam), 1:200 mouse M2 anti-FLAG, 1:500 rabbit anti-HA and 1:1000 Streptavidin-Alexa-647 direct conjugate (Invitrogen).

2.4 Molecular Methods

2.4.1 RNA extractions

All RNA extractions were carried out using the RNAqueous-Micro Kit (Life Technologies, cat # AM1931) and included DNAseI treatment as per the kit instructions.

2.4.2 cDNA generation

5μg of isolated RNA was used for a two-step reverse transcription reaction. Step one: 5μg of isolated RNA + 1μL 50uM oligo dT primers + 1μL 10mM dNTP mix in a 13μL volume of H_2O. This annealing mix was heated to 65°C for 5 minutes and incubated on ice while the following mixture was added: 4μL of 5x First strand synthesis buffer (Invitrogen), 1μL of 0.1M DTT, 1μL of RNAseOUT Recombinant RNase inhibitor (Invitrogen) and 1μL Superscript III reverse transcriptase (Invitrogen). All reverse transcription experimental samples were accompanied by a reverse transcriptase negative control sample. Both reactions were incubated at 25°C for 5 minutes, then at 50°C for 1 hour.

2.4.3 RT-QPCR analysis

All RT-QPCRs were run on the Applied Biosystems 7500 Fast Real-Time PCR System using Fast SYBR Green (ThermoFisher, cat #4385612) reagent in MicroAmp fast Optical 96-well plates (ThermoFisher, cat #4346906). Primers were designed to amplify 100bp fragments and ordered from IDT. Standard RT-QPCR protocols were used (see Appendix 2 for primer details).

2.4.4 NanoString gene expression profiling

NanoString technology was used to profile different experimental conditions on the expression levels of a large number of genes simultaneously from low starting material. All Nanostring experiments were done on material lysed in the lysis buffer of the RNAqueous-Micro Kit (Life Technologies, cat # AM1931), mixed with the NC-GRN custom capture probe codeset and barcodes, and loaded into the Nanostring nCounter system. The NanoString nCounter platform captures individual transcripts using barcoded probes, which are then digitally counted subsequent to microscopic imaging. Each target probe has a specific barcode, accordingly the data generated is a count of the numbers of each target transcript within the tissue provided. This is done two robotic platforms, one responsible for the chemistry, the other for the digital counting. The target genes are user-defined; a NC specific probe-set was used in this work (see Appendix 4 for probe-set details). This method involves two 50bp probes that capture their target mRNA; where the reporter probe carries the barcode and the capture probe allows the complex to be immobilised. Once the excess probes are washed off, the probe/target complexes are immobilised and aligned in an nCounter cartridge. The cartridge is then imaged, and the data counted in the Digital Analyser robot.

2.4.5 Biotagging-compatible construct cloning

All biotagging compatible constructs contain the Avi-TeV-FLAG tag cassette. All tagged factors were driven by either tissue-specific elements or ubiquitously by the chicken pCAGGS promoter as indicated in the text. All vectors carried ampicillin resistance genes for selection purposes.

The In-Fusion HD kit (Clontech, cat #638910) was used as an efficient and ligation-independent cloning method for certain constructs, as indicated below. It involves the single-step, seamless and directional integration of a given insert into a target vector, subsequent to target sequence amplification using In-Fusion compatible

primers.

2.4.6 Avi-FoxD3-EGFP construct cloning

The final Avi-FoxD3-EGFP construct contained an N-terminal Avi-tagged FoxD3 and a C-terminal EGFP fluorescent reporter. The Avi-FoxD3-EGFP construct was generated by restriction cloning of a fusion PCR fragment into the original Avi-FoxD3 construct. This original Avi-FoxD3 construct contained the Avi-tag cassette at the N-terminal of FoxD3 only and was previously cloned within the lab (original clone PBS cFoxD3 #217 (identified by C.Erickson and previously used by T.Sauka-Spengler). The C-terminal portion of FoxD3 was amplified from Avi-FoxD3 using a GC-rich PCR system (Roche, cat #12140306001) and fusion PCR compatible primers. This C-terminal FoxD3 fragment was fused to a 2A-EGFP sequence using a fusion PCR protocol, and subsequently was restriction cloned into the Avi-FoxD3 vector using PvuII and XbaI.

2.4.7 Citrine-FoxD3-Avi construct cloning

The final Citrine-FoxD3-Avi construct contained a C-terminally Avi-tagged FoxD3, and an N-terminal Citrine fluorescent reporter. The FoxD3 ORF was amplified from the original cFoxD3 clone (identified by C.Erickson and previously used by T.Sauka-Spengler) using In-Fusion compatible primers. The In-Fusion compatible fragment was cloned using the In-Fusion HD kit (Clontech, cat #638910) into a digested vector containing the Citrine reporter and purification cassette (Citrine-2a-//-FLAG-TeV-Avi). The Citrine-2a-GAL4DBD-FLAG-TeV-Avi vector was prepared by digestion with PvuII and SmaI to create the In-Fusion kit ready Citrine-2a-//-FLAG-TeV-Avi linearised vector.

2.4.8 FoxD3-Avi construct cloning

The final FoxD3-Avi construct contained a C-terminally Avi-tagged FoxD3 only. The Citrine-FoxD3-Avi construct was used as a template to amplify the tagged FoxD3 region for the FoxD3-Avi construct. The Citrine-FoxD3-Avi construct was restriction digested by NotI and NheI to release the tagged insert and generate a linearised vector. The FoxD3-Avi containing fragment was then digested using NheI and ligated back into the linearised vector.

2.4.9 Avi-Sox10 construct cloning

The final Avi-Sox10 construct contained an N-terminally Avi-tagged Sox10 only. Avi- Sox10 was cloned using In-Fusion HD Cloning (Clontech, cat #638910). The Avi-TeV- FLAG cassette was amplified from the Avi-FoxD3 construct using primers with InFusion compatible tails.The 1.4kb Sox10 ORF was amplified from the PBS Sox10 cDNA clone (previously used in lab by T. Sauka-Spengler) using primers with InFusion compatible tails. A novel 2.4kb Sox10 enhancer (R.Williams, unpublished data) was cloned into an enhancer activity testing vector containing a Citrine reporter using GoldenGate cloning. This vector was previously digested by NcoI and XbaI restriction enzymes and had the Avi cassette and Sox10 ORF cloned into it using the In-Fusion cloning kit.

2.4.10 Biotin-ChIP-QPCR

Biotin-ChIP-QPCR was conducted to assay for increased representation of sites bound by Avi-tagged TFs compared to input material. 8-10 dorsal cranial regions were dissected and collected at 10-12ss chicken embryos previously electroporated with biotagging constructs. This stage was chosen to ensure high levels of FoxD3 and Sox10 expression respectively. The tissue was resuspended in a nuclei expulsion buffer (0.5% NP-40, 0.25% TritonX-100, 10mM Tris-HCl pH7.5, 3mM CaCl2, 0.25M sucrose, 1mM DTT, 0.2mM PMSF, 1x protease inhibitors in water), then physically

homogenised into a single-cell suspension using a dounce homogeniser to release the nucleii into suspension. Samples were then cross-linked (1% formaldehyde for 8 minutes), lysis (0.6% SDS, 50mM Tris pH 8.0, 10mM ETDA & 1x protease inhibitors) and sonicated (approximate size 400-800bp). Biotin-ChIP was performed as per usual protocols (Betancur, Bronner-Fraser, & Sauka-Spengler, 2010b; He & Pu, 2001; Trinh et al., 2017) . All samples were pre-cleared using 50μl protein A beads (Thermo Fisher cat #10006D) prior to addition to 50μL Streptavidin MyOne T1 beads (Invitrogen, cat #65601) blocked in 1% biotin-free BSA in PBS per sample. Input samples were collected subsequent to preclearing and prior to addition of streptavidin MyOne T1 beads. Samples were washed twice using SDS wash buffer (2% SDS , 10mM Tris pH8, 1mM EDTA), and washed 8 times using RIPA wash (50mM Hepes-KOH pH8, 500mM LiCl, 1mM EDTA, 1% NP-40, 0.7% NA-Deoxycholate, 1x protease inhibitors in water). The last wash conditions are in 1x TE with 50mM NaCl. After the last wash was removed, both input and experimental samples were un-crosslinked at 65°C in a thermomixer at 1300rpm in SDS elution buffer (50mM Tris-HCl pH8, 10mM EDTA, 1% SDS in water). Biotin-ChIP material was then RNAse and proteinase treated prior to phenol chloroform precipitation to purify the isolated DNA. Material was then used as a template for QPCR (see Appendix 2 for details).

2.5 Genome- and Epigenome-Editing Toolkit Methods

2.5.1 sgRNA cloning

All synthetic guide RNAs (sgRNAs) were generated by annealing forward and reverse oligos using the following protocol: 94°C 5 minutes, 64°C-34°C (with a temperature decline of 1°C per minute) for 30 minutes. Annealed sgRNA oligos were cloned into U6 mini vectors using a GoldenGate BsmBI reaction: 75ng U6 vector, 4ng sgRNA, 1μL BsmBI (NEB cat #R0580), 2μL 10x T4 Ligase Buffer, 1μL T4 Ligase (NEB cat# M0202S) in a 20μL H_2O reaction mixture. The GoldenGate reaction was

as follows: 37°C 5 minutes, 16°C for 10 minutes, 10 cycles of: 50°C 5 minutes, 80°C 5 minutes. 5μL of these reactions were transformed into competent cells and selected on tetracycline plates. Single colonies were screened for successful integration by overnight culture, miniprep and sequencing with the U6_seq primer (see Appendix 3 for details).

2.5.2 pX330-Cas9 construct

The active pX330-Cas9 construct was previously cloned within the lab (Williams et al., 2018).

2.5.3 pX330-dCas9-KRAB construct cloning

This construct contains a catalytically dead Cas9 fused C-terminally to a KRAB domain. A synthetic DNA fragment containing the KRAB sequence was cloned using the In- Fusion HD Cloning kit (Clontech, cat #638910) into the EcoRI linearised pX330-dCas9- LSD vector (previously cloned within the lab, Williams et al, 2018).

2.5.4 dCas9-KRAB construct cloning

These are modifications of constructs kindly shared by the Gerschbach group (Thakore et al., 2015).

2.5.5 pCI-dCas9-KRAB-GFP

This construct contains a GFP fluorescent reporter at the C-terminal end of a catalytically dead Cas9 fused C-terminally to a KRAB domain. The pCI-dCas9-KRAB-T2A-GFP construct was generated by amplifying the 5.3kb dCas9-KRAB-T2A-GFP sequence from pLV hUbC dCas9 KRAB T2A GFP (kindly donated by P.Thakore) and cloning into the pCI-H2B-RFP vector linearised by NotI and XhoI. Both vectors contain promoters that are ubiquitously expressed in chicken embryos.

2.5.6 High Resolution Melt Analysis (HRMA) analysis

HRMA analysis reports differences in melting temperature between PCR amplicons in a qPCR melt-curve like assay. In this application, CRISPR-Cas9 induced indels are predicted to change the melting profiles of fragments spanning the target site. Thus, by comparison with amplicons from Cas9 only or Cas9 + scrambled sgRNA samples, effective alterations from the wild-type profile can be seen.

Genomic DNA (gDNA) was extracted from single embryos using the Purelink Genomic DNA mini kit (Invitrogen). Primers were designed to generate a 100bp product spanning the gRNA cut site. HRMA PCR was performed using Hotshot Diamond PCR mastermix (Client Lifescience, #HS002-TS) together with LC Green Plus dye (BioFire Diagnostics, #BCHM-ASY-0005). Reaction solutions were cycled on a C1000 TouchTM Bio-Rad thermal cycler (95°C - 5mins, 49 cycles: 95°C - 20 secs, 60°C Ð 20 secs, 72°C - 30 secs, followed by 72 °C - 2 mins, 95 °C - 30 secs, 25 °C - 30 secs, 95°C - 30 secs). A final melt curve (65 °C - 95 °C with increments of 0.2 °C for 0.02 secs) was included for melt analysis. BioRad Precision Melt analysis software was used to visualise and analyse the data.

2.5.7 T7 Endonuclease Assay

The T7 Endonuclease Assay was used to determine genome targeting efficiency by CRISPR-Cas9. It was conducted upon heated and cooled 1kb PCR fragments amplified from genomic DNA. Differences between T7 endonuclease digestion of amplicons from CRISPR-targeted and control material reflects changes to the DNA, attributed in this context to the activity of targeted CRISPR-Cas9. The endonuclease recognises non-matched DNA, which is revealed by the heating and cooling of a pool of material with differential and independent editing events.

Genomic DNA was extracted from single embryos using the Purelink Genomic DNA mini kit (Invitrogen). Primers designed to generate a 1kb product spanning

the gRNA cut site were used to generate a PCR fragment. This fragment was gel purified using the Wizard SV Gel and PCR Clean-Up System (Promega, cat # A9282). 200ng of purified PCR product was incubated with 2μL NEB Buffer 2 in a final volume of 19μL at 95°C for 5 minutes. The temperature was then reduced at the rate of -2°C per second from 95-85°C, and -0.1°C per second till 25°C was reached. Subsequently, 10 units of T7 endonuclease I (NEB, cat #M0302) were added to the cooled mixture and incubated at 37°C for 15 minutes. The reaction was halted by the addition of 2μL of 0.25M EDTA and visualised immediately on a 1.5% agarose gel.

2.6 *In vitro* Protein Methods

2.6.1 Protein extraction from chicken embryos

To extract protein from chicken embryos for downstream use in western blots (see below) embryos were dissected from extra-embryonic material, washed in Ringer's solution containing 1x protease inhibitors. Embryos were spun-down in a table-top centrifuge at 4°C at 16,500g, and all solutions removed. Samples were then incubated on ice in 5 volumes of 1x buffer H (0.2mM HEPES pH 8.0, 0.015mM MgCl2, 0.1mM KCl, 1x protease inhibitors, 1mM DTT, 0.4mM PMSF) for 15 minutes (approx. 5μL per embryo in tube). Embryos were homogenised on ice using a 28.5-gauge needle pre- coated with buffer H. Subsequent to another 15-minute incubation on ice, the samples were centrifuged for 5 minutes at 4°C at 16,500g. This supernatant is the cytoplasmic fraction and was stored at -80°C prior to use. The pellet was resuspended in an equal volume of 1x buffer D (1.5mM MgCl2, 200mM HEPES pH 8.0, 0.2mM EDTA, 420mM KCl, 20% Glycerol, 1x protease inhibitors, 1mM DTT, 0.4mM PMSF) as was used as buffer H in the first step. Samples were nutated at 4°C for 30 minutes prior to centrifugation at 16500g for 5 minutes at 4°C. This supernatant is the nuclear fraction and was separated and stored at -80°C for prior to

use. Total protein concentration was determined using the Pierce 660nm protein assay kit (Sigma, cat #22660), subsequent to mixing equal volume of cytoplasmic and nuclear protein fractions.

2.6.2 Western Blotting

To confirm presence and absence of various proteins, proteins were extracted (as detailed above) and equal quantities of protein per sample were run on 15% precast Tris-glycine polyacrylamide gels (BioRad, cat #4561083) under SDS denaturing conditions (0.01% SDS in Tris-glycine). The BioRad Trans-Blot Turbo Transfer system was used to transfer proteins onto nitrocellulose membranes. The membranes were blocked in the appropriate blocking solution 3 x 20 minutes. All blots detected with antibodies were blocked using normal blocking solution (150mM NaCl, 5% Non-fat dried milk, 50mM Tris pH 8.0, 0.1% Tween-20). As milk contains biotin, Strep-HRP blots were blocked using biotin-free streptavidin block (5% BSA (Biotin-free), 0.2% Tween-20, 20mM Tris pH 8.0, 137nM NaCl). Primary antibodies were diluted in respective blocking solutions and incubated overnight at 4°C. Excess antibody was washed away using blocking solution or streptavidin wash solution respectively (6x 20 minutes). Subsequent to washes, secondary HRP-conjugates (GE-Healthcare) were added at 1:2000 dilution in block for 2 hours with agitation at room temperature. Biotinylation was detected using a direct streptavidin-HRP conjugate, therefore these blots were developed after washes removed excess conjugate. Signal was developed using the ECL-detection kit from (GE-Healthcare, cat #RPN2232) following kit instructions.

Primary antibodies: 1:2000 rabbit anti-GFP (Torrey Pines, cat #TP401), 1:1000 mouse anti-FLAG (Sigma, cat#F1804), rabbit anti-HA (Abcam, ab9110), 1:20,000 mouse anti-GAPDH and Streptavidin-HRP (Invitrogen).

Chapter 3

Establishing an Advanced Toolkit for Exploring Gene Regulatory Interactions using the Chicken Embryo Model System

3.1 Introduction

Understanding the potential cranial NC specific dysregulation of CHD7 in CHARGE syndrome, requires placing CHD7 into the broader context of the GRN known to drive cranial NCC specification, behaviour and differentiation. GRNs are schematic representations of direct and indirect gene product interactions driving developmental processes. To place CHD7 in this wider cranial NC-GRN, the potential regulatory sites and upstream NC- specific regulators acting via those sites, must be identified.

Much of the NC-GRN has been elucidated within the chicken model system (reviewed by Meulemans and Bronner-Fraser (2004); Nikitina et al. (2009); Sauka-Spengler and Bronner-Fraser (2008); Simões-Costa and Bronner (2015)). Comparative genome biology has classically been used to identify conserved regulatory regions, which often contain enhancers (reviewed by Furlong and Levine (2018)).

Chicken embryos are particularly amenable to physical manipulation of the NC (reviewed by Le Douarin (2004); Stern (2005)). Consequently, putative conserved regulatory elements can be tested by introduction of fluorescent reporter constructs, via electroporation, into the chicken embryo (reviewed by Uchikawa (2008)). Similarly, morpholino electroporation to interfere with genes has been a hugely useful tool to test elements of the cranial NC-GRN. For example, morpholinos were used to demonstrated the importance of FoxD3 in NC lineage establishment in the chicken model system (Kos et al., 2001).

While embryological techniques have always been easily employed in the chicken system, certain techniques have proven to be difficult to employ in this model. During early chicken development, the cranial NC is a small cell population. Techniques such as chromatin precipitation (ChIP), used to demonstrate direct interactions in a GRN, normally requiring millions of cells have been unfeasible in this context. To ameliorate this low cell number limitation, and to allow assessment of direct TF binding in the cranial NC, the recently developed biotagging technique could translated for use in the chicken (Trinh et al., 2017). Biotagging technology, relies on the tagging of factors of interest and their subsequent biotinylation by a biotin ligase. This biotinylated factor can be extracted with high efficiency from its endogenous context using streptavidin binding. The translation of this technology into the chicken system would allow for profiling of key NC TF binding both globally, and to putative regulatory sequences. Furthermore, while morpholino based loss-of-function experiments have been a standard for many years, developments in genome engineering have opened the door to more precise and direct gene editing via the CRISPR-Cas9 system. Only a few recent attmepts to utilize existing CRISPR-Cas9 constructs in the chicken embryo have been made with limited success (Veron et al., 2015; Shashank Gandhi and Bronner, 2016).

Here I present our work establishing an advanced toolkit, composed of biotagging TFs and CRISPR-Cas9 based genome engineering tools for the chicken model system (Williams et al., 2018), which I required to establish the position of CHD7 in

the cranial NC- GRN.

3.1.1 Resolving Gene Regulatory Networks

Developmental processes are governed by the coordinated actions of large sets of genes, expressed in precise spatio-temporal patterns. Gene regulatory networks (GRNs) define the specific causal links between these developmental processes, and the regulatory elements underlying them. In particular, GRNs describe the role that each gene product plays in activating or inhibiting downstream gene expression. GRNs detail whether the interactions are known to be direct or indirect, and whether they require the coordinated actions of many gene products (Davidson et al., 2003). Elucidating these regulatory interactions is crucial to our improved understanding of the molecular mechanisms underlying human development and disease.

This precise transcriptional spatio-temporal control is achieved by the combined activity of transcriptional regulators, known as inputs, upon *cis*-regulatory sequences such as enhancers, silencers and insulators (Davidson and Levine, 2008; Davidson et al., 2003). Enhancers serve to promote downstream gene transcription by facilitating upstream transcription factor (TF) mediated recruitment of RNA polymerase II to core promoters/ proximal regulatory elements at transcription start sites (Khan et al., 2013). As such, enhancers define the architecture of GRNs by regulating the relationship between upstream TF inputs and downstream gene output (Davidson & Levine, 2008). Silencers are the opposite of enhancers; the factors bound to this sequence repress downstream gene transcription (Kolovos et al., 2012).

3.1.2 Tools currently available for studying enhancers

Antibody-based Chromatin Immunoprecipitation (ChIP) of histone modifications known to associate with enhancers can be used to identify putative enhancer elements (Shlyueva et al., 2014). For example, histone 3 lysine 4 mono-methylation

(H3K4me1) associates with active enhancers, whilst histone 3 lysine 27 acetylation (H3K27Ac) associates with active promoters and enhancers (Heintzman et al., 2007) (recently reviewed by Calo and Wysocka (2018)). While antibody-mediated ChIP-sequencing of these histone modifications facilitates identification and verification of putative enhancers, they require a significant amount of starting material.

To date, many putative enhancers have been identified by comparative genome analysis of highly conserved non-coding DNA sequences and DNAse I hypersensitivity assays (Bellard et al., 1986; Khan et al., 2013; Shlyueva et al., 2014). However, the recent development of Assay for Transposase-Accessible Chromatin with high throughput sequencing (ATAC-seq) has vastly improved identification of active enhancers. ATAC- seq reveals sites of accessible chromatin, much like DNAseI hypersensitivity, whilst requiring much less starting material than previous assays (Buenrostro et al., 2013). Finally, using next generation Capture-C protocols it is also possible to establish the topologically associated domain surrounding an enhancer, and locate the promoters most likely to be regulated by it via mapping of physical interaction sites (Davies et al., 2016).

Given GRNs require information on specific interactions between inputs and downstream outputs, establishing the genomic location of enhancers is the first step towards identifying the TFs that bind these regulatory elements, and the downstream consequences of these binding events. Hence profiling TF binding at putative *cis*- regulatory regions is a vital step in the resolution of a GRN.

Antibody-mediated ChIP-sequencing is a well-established method for profiling TF binding sites genome-wide, and identifying enhancers and promoters bound by specific TFs (Nakato and Shirahige, 2017). Traditional ChIP protocols require the generation of protein-specific antibodies, which can have variable affinity and thus require high starting material, which is a limiting factor when studying small cell populations in development (Kim et al., 2009). Whilst ChIP-sequencing helps locate enhancers and promoters bound by the TF in question, it does not indicate the nature of the regulation exerted. Classically, morpholino or RNA interference

(RNAi) based TF knockdown coupled with expression profiling is used to determine the effect of TF knockdown on downstream gene expression (Bill et al., 2009; Corey and Abrams, 2001; Stainier et al., 2017; Stern, 2005).

Once identified, putative enhancers can be tested using fluorescent reporter assays *in vivo*, a particular speciality of the chicken model system (Uchikawa et al., 2003; Itasaki et al., 1999; Muramatsu et al., 1997, 1998, 1996) (reviewed by Stern (2005); Uchikawa et al. (2017)). For example, the chicken model system can be used to transiently express three spectrally distinct fluorescent reporters, allowing the characterisation of specific subpopulations based on combinations of enhancers expressed (Williams et al., 2018). This type of reporter assay in chicken also lends itself to dissecting upstream regulatory interactions, by allowing the perturbation of putative TF binding sites within the region in question (Betancur et al., 2011, 2010b; Simões-Costa et al., 2012).

3.1.3 The chicken model system

First used in the 19th century at the advent of developmental biology, the chicken embryo is now an established model to study early embryonic development. Chicken embryos share morphological complexity and general embryonic development with mammalian embryos, including humans. Chicken embryos are easily sourced, cultured *ex-ovo* and highly amenable to microsurgical manipulation experiments (Reviewed by Stern (2005); Uchikawa et al. (2017). Development of transplantation based quail-chicken chimera methods, which enabled tracking of NCCs during migration and differentiation, led to a comprehensive description of the origin and fate of this cell population (Le Douarin, 1973; Le Douarin and Teillet, 1973; Balaban et al., 1988; Teillet et al., 1999; Agrawal and Lwigale, 2015). Additionally, the chicken genome is very compact, low chromosomal translocation rate and clear synteny between chicken and other amniote vertebrates, including mammals (Burt, 2005). There is a high level of sequence conservation in intronic and flanking noncoding

regions likely to contain important regulatory elements. This makes it a valuable tool for the identification of putative regulatory regions such as enhancers (Burt, 2005; Stern, 2005; Brown et al., 2003; Khan et al., 2013; Uchikawa et al., 2003).

However, identifying enhancers by conservation is limited as it does not enable identification of species-specific elements, nor does it account for functional synteny in the case of elements which have different sequences. TF binding sites can also change and consequently the sequences of the enhancer though which they function presumably also evolve. With the advent of genomic analysis and transcriptomics, enhancers can be identified using more advanced bioinformatic approaches, and subsequently validated and tested with ease using established chicken embryological methodologies (Betancur et al., 2010a; Simões-Costa et al., 2012).

3.1.4 Using the chicken system to resolve the NC-GRN

The chicken model is an established system for studying NC development. Quail-chicken chimera studies have shown that the cranial NC is the most multipotent of the NC axial divisions as it contributes the widest range of these derivatives including facial cartilage and bone, teeth and cranial ganglia (Le Douarin and Kalcheim, 1999; Le Douarin, 2004). Aberrations in the cranial NC are implicated in a multitude of human disorders such as Di George and CHARGE syndromes (reviewed by Trainor (2010); Snider and Mishina (2014)), in addition to cancers such as neuroblastoma and melanoma (reviewed by Jiang et al. (2011); Menendez et al. (2013); Kulesa et al. (2013); Maguire et al. (2015)). NC ontogeny is controlled by a multimodule gene regulatory network (NC-GRN), which can be organised into hierarchical sub-networks reflecting the different stages of NC formation (Sauka-Spengler and Bronner-Fraser, 2008; Sauka-Spengler and Barembaum, 2008; Meulemans and Bronner-Fraser, 2005; Nikitina et al., 2009; Simões-Costa and Bronner, 2015). Identifying active *cis*-regulatory elements to important NC genes at key times of NC ontogeny, in addition to the upstream TFs regulating them, is an important step

towards the resolution of the NC-GRN.

3.1.5 Current tools used to dissect the NC-GRN in chicken embryos

Traditionally, gene expression patterns in chicken are investigated using immunocytochemistry and *in situ* hybridisation protocols. Fluorescent *in situ* hybridisation techniques, e.g. hybridisation chain reaction (HCR), have increased the number of transcripts being probed within a single embryo, whilst improving the subcellular resolution (Choi et al., 2016; Dirks and Pierce, 2004). Changes in these patterns are used to assess the effect of TF knockdown *in vivo*, further revealing the hierarchy of genes in a GRN in question. Traditionally, experimental probing of the NC- GRN was mainly restricted to loss- and gain-of-function involving cDNA, mRNA, morpholino or antisense oligonucleotide injections (Sauka-Spengler and Bronner-Fraser, 2008).

The development of electroporation methodologies facilitated the widespread use of the chicken model system to identify and characterise new regulatory regions of DNA (see Figure 3.1.) (Nishi et al., 1996; Muramatsu et al., 1997; Itasaki et al., 1999; Uchikawa et al., 2004; Timmer et al., 2001). Electroporation involves the transfer of plasmid DNA into living cells through the introduction of nanometre pores into cell membranes by the application of electric pulses (Shafritz, 1986). Electroporation facilitates a swift read-out of transgene activity in chicken embryos, which is a useful characteristic for studying developmental enhancers (Kvon, 2015).

To test enhancer activity using chicken embryos, putative enhancer elements are cloned into reporter vectors containing a minimal promoter upstream of a fluorescent gene and electroporated into the chicken at various stages of development (Itasaki et al., 1999; Muramatsu et al., 1997, 1998, 1996; Simões-Costa et al., 2012). The coupling of enhancer reporter plasmid electroporation with *ex-ovo* chicken embryo culture allows frequent monitoring for enhancer reporter activity *in vivo*. As

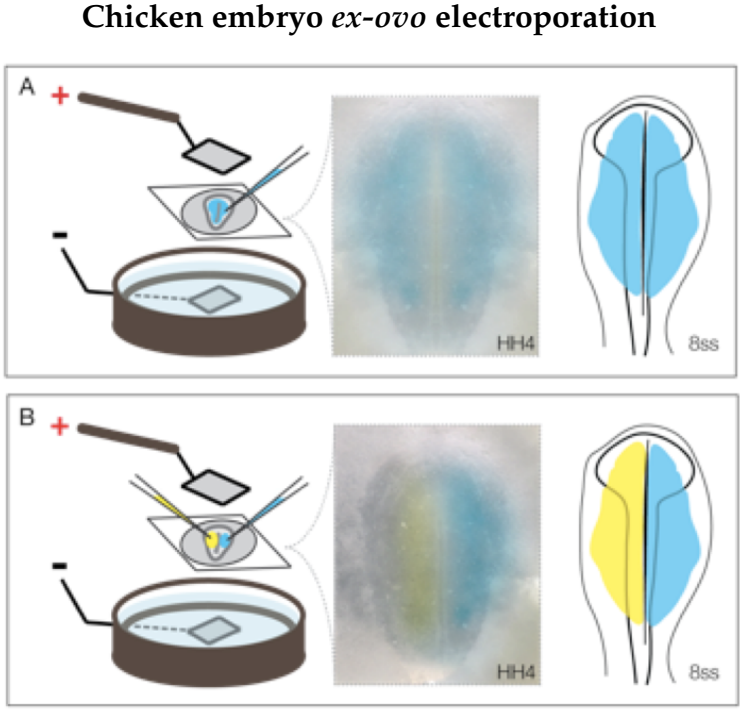


FIGURE 3.1: (A) whole embryo electroporation is conducted by injecting plasmid (blue) into the sac beneath the epiblast at HH4 and applying an electric current across the embryo. The inset shows a view of an embryo down the eyepiece, prior to electroporation, and a schematic of the expression of the electroporated plasmid later in development, in the cranial region of an 8ss embryo. (B) For bilateral electroporation, a different plasmid is injected on either side of the embryonic midline prior to application of electric pulses. The distribution of the plasmids can be seen in the inset image, which shows a view of an embryo down the eyepiece prior to electroporation. Different colours of vegetable dye (yellow and blue) are used to mark each plasmid mixture. As shown in the schematic, this separation is maintained as the embryo grows, as the expression of each plasmid is confined to one side of the embryo. This technique allows the use of one side of the embryo as in internal control experiment.

NCCs do not cross-the midline bilateral electroporation techniques, where different constructs are electroporated on opposite sides of the midline, can be employed (see Figure 3.1. panel B) (Hu et al., 2012; Simões-Costa et al., 2012). The use of bilateral electroporation permits experimental and control conditions to be tested within the same embryo. These enhancer-reporter vectors can also be multiplexed into multi-colour assays to characterise populations by differential enhancer activity (Seo et al., 2010). Multiplexing enhancer reporter constructs allows the separation of sub- populations of cells which may employ an enhancer for a common factor, but differentially employ other enhancers for other factors.

3.1.6 Expanding the chicken toolkit

While these tools are able to reveal preliminary regulatory information of the NC-GRN, they do not elucidate finer regulatory details such as intermediary factors or direct TF- enhancer-promoter interactions (Meulemans and Bronner-Fraser, 2004). We are particularly interested in identifying and characterising NC specific enhancers active in the cranial region of the chicken embryo at HH7 - HH10 as this is a key stage in NC ontogeny- the characteristic EMT (Hamburger and Hamilton, 1951). The NC is easily observable at this stage using fluorescent reporters, as it is a small cell population with approximately 500-1000 cells per cranial region (Simões-Costa et al., 2012). To investigate the nature and direct association of upstream TFs with identified enhancer elements we need to perform chromatin immunoprecipitation (ChIP) assays. However, performing antibody-based TF ChIP-sequencing presents several technical challenges. Firstly, current protocols dictate the need to use at least one million cells per pull-down, a feat easily achievable using *in vitro* methods but not as yet *in vivo* in the chicken embryo (Kim et al., 2009). Secondly, ChIP-grade antibodies are not available for all the TFs of interest. Additionally, antibody-antigen interactions are not always strong enough to sustain the stringent washes necessary to eliminate ChIP background and enable significant amplification of pull-down material in conditions of small cell numbers. Thus, I adapted a new method to perform tagged-TF ChIP in the chicken embryo. In the following chapter the use of *in vivo* biotinylation, and optimisation of biotin-ChIP procedures in the chicken embryo are presented.

Another improvement to the toolkit for studying TF/enhancer interactions consists of implementing highly efficient transient genome engineering-mediated knockdown in the chicken embryo. Traditional gene knockdown methods required to characterise the regulation of TFs on the identified enhancers are known to have variable levels of knockdown and have off-target effects. The advent of CRISPR-Cas9 technology promises to bypass many of these issues when applied to the chicken

system.

3.2 The 'Biotagging' Principle and Toolkit for Chicken

To profile global binding of key NC TFs during NC development in the chicken, I adapted the Sauka-Spengler zebrafish 'biotagging' system. This system couples streptavidin- biotin affinity purification of Avi-tagged TFs with next generation sequencing to reveal the global binding profiles of these TFs (see Figure 3.2.).

'Biotagging' is a genetically-encoded binary approach to enable regulatory profiling and characterisation of molecular interactions in zebrafish embryos, developed in our lab. It utilises *in vivo* biotinylation as previously described in other model organisms (de Boer et al., 2003; Deal and Henikoff, 2011). By co-expressing a biotin acceptor peptide (Avi-) tagged TF-of-Interest and a bacterial biotin ligase (BirA) within the same cell(s), the addition of a single biotin residue by BirA to the Avi-tagged TF-of-interest allows the use of downstream streptavidin-based affinity purification methods to isolate and characterise DNA elements bound by the TF-of-interest (see Figure 3.2.) (Beckett et al., 1999; de Boer et al., 2003; Kim et al., 2009; Schatz, 1993). BirA transfers the biotin moiety from its site on the endogenous biotin carboxyl carrier protein (Choi-Rhee and Cronan, 2003). Furthermore, this novel technology uses tissue-specific elements to drive expression of the biotagging components *in vivo*. Using tissue specific enhancers allows the precise extraction of information from user-defined cell populations. In addition, it ensures physiological levels of Avi-tagged TF expression, avoids overexpression and ectopic expression effects, as enhancing strength is determined by the endogenous factors driving expression via the particular enhancer element.

Biotin labels are small and stable, and previous work indicates that biotinylation of Avi- tagged proteins does not alter protein function, interactions or DNA binding (de Boer et al., 2003) therefore making it suited to be utilised in studies dissecting eukaryotic biological processes (Beckett et al., 1999; He and Pu, 2001; Kim et al., 2009).

The streptavidin-biotin interaction is one of the strongest known non-covalent interactions in nature (Kd = 10-15M) and these bonds are unaffected by extremes of pH, organic solvents, denaturing agents or temperature (Holmberg et al., 2005).

Subsequent to isolation from crude nuclear extracts via direct binding to streptavidin beads, the strength of the biotin-streptavidin bond allows very high stringency washing conditions. This permits efficient purification of biotinylated proteins with the associated proteins and DNA from small cell populations, whilst simultaneously reducing the background signal usually observed in antibody-ChIP (see Figure 3.2. Panel B) (Kim et al., 2009; Kolodziej et al., 2009; Rudra et al., 2012). Pairing *in vivo* biotinylation with NGS technologies provides a powerful tool to study TF binding at a systems level, revealing global patterns and novel biological insights that cannot be acquired via traditional methods alone.

3.2.1 Biotagging key NC TFs FoxD3 and Sox10 for use in the chicken model system

To test Avi-tag accessibility for *in vivo* biotinylation, I generated both N-terminal and C- terminally Avi-tagged FoxD3 constructs (see Figure 3.3.). Constructs containing additional fluorescence reporters, separated upon translation by ribosome-skipping 2A peptide sequences, were cloned to confirm tissue specific expression of the Avi-tagged TFs without necessitating electroporation of additional plasmid DNA constructs. The BirA construct contains a triple HA tag in addition to a mCherry reporter separated by ribosome-skipping 2A peptide. HA tag serves to confirm localisation of the BirA enzyme (nuclear) and the mCherry fluorescent protein allows identification of BirA-expressing cells after electroporation. To ensure equal co-transcriptional representation of the Avi- TF and the fluorescent reporter, the 'self-cleaving' 2A signal was used. The 2A sequence induces ribosomal skipping of the glycyl-prolyl peptide bond at its C terminus (Donnelly et al., 2001). The outcome of this skipping is that a single ribosome translates both the sequence before the

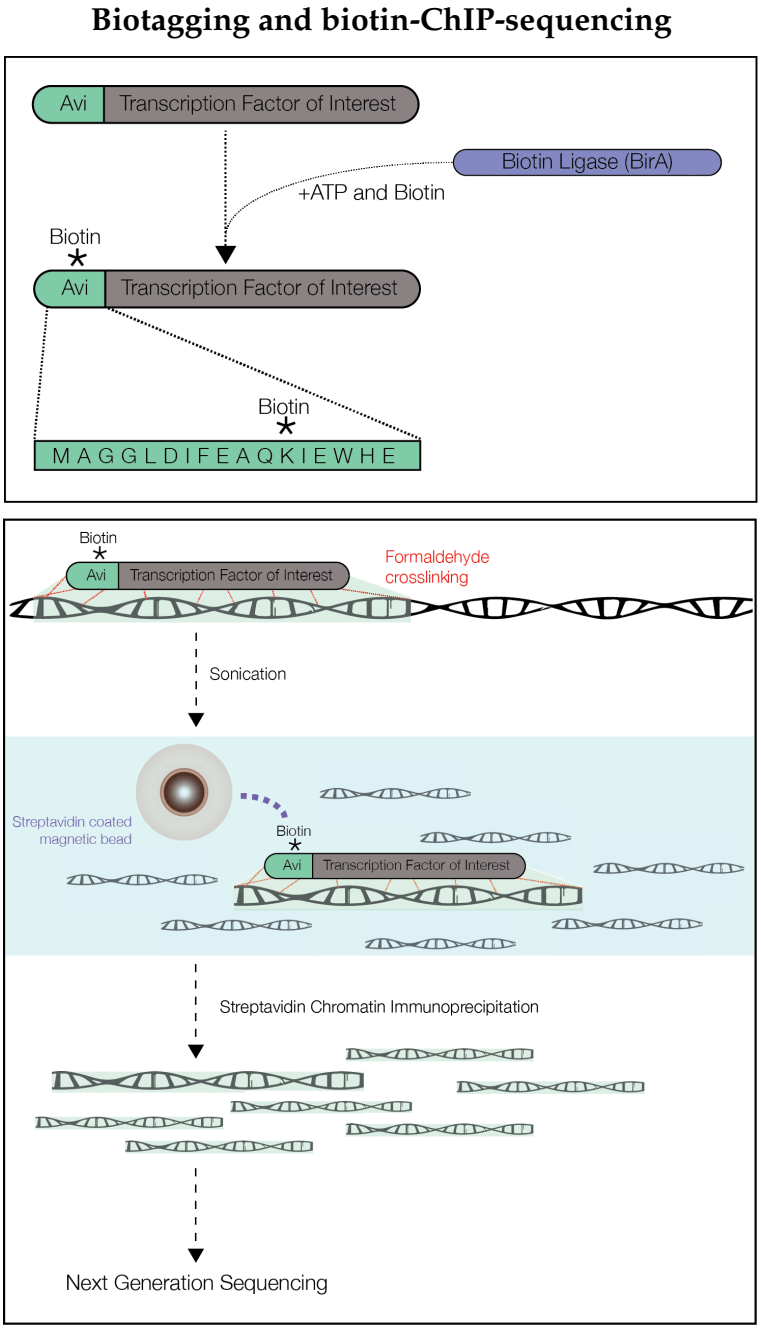


FIGURE 3.2: (A) A TF of interest is tagged with a 17 amino acid Avi-tag peptide. The biotin ligase, BirA, uses ATP to catalyse the addition of a single biotin residue to a lysine in the Avi- tag. (B) The Avi-tagged TF is crosslinked to its associated DNA using formaldehyde. After sonication of the crosslinked DNA-protein complex, streptavidin coated magnetic beads are used to specifically purify the DNA fragments that are associated with the Avi-tagged TF. The purified DNA fragments can then by sequenced via next generation sequencing or assayed by qPCR.

2S sequence and that after it. This results in equal numbers of separated proteins, thereby circumventing potential alterations of function by a fusion reporter protein.

The Avi-tagged NC TFs FoxD3 and Sox10 were placed under the control of known NC-specific enhancers to ensure endogenous levels of expression (see Figure 3.3.). FoxD3 was driven by the cranial NC specific FoxD3 enhancer 'NC1' if premigratory cranial NCCs were targeted. Alternatively, by the mutated NC1 enhancer ('M3') if migrating NCCs were also to be captured (Simões-Costa et al., 2012). The basal chicken B-actin promoter (pCAGGS) was used to drive high and ubiquitous expression necessary to acquire sufficient material for protein validation by western blot. The Avi-tagged Sox10 was driven by a novel NC specific Sox10 enhancer identified in the lab (R.Williams & T. Sauka-Spengler, unpublished data).

The biotinylatable Avi tag cassette was cloned at the N- and C- terminal of FoxD3, and the N-terminus of Sox10. The Avi-tagged FoxD3 constructs were generated both with and without fluorescent reporters. In the constructs containing fluorescent reporters, a 2A ribosomal skipping peptide sequence was used to separate the reporter transcript from that of the Avi-tagged TF; ensuring an approximate 1-to-1 ratio of fluorophore to Avi-TF. In constructs containing the reporter, the respective 2A and fluorophore sequences were at the opposite terminus to the Avi-tag.

3.2.2 Characterising Avi-tagged TF constructs expression and confirmation of Avi-tagged FoxD3 protein *in vivo* biotinylation

To initially confirm expression of BirA and Avi-tagged FoxD3 protein *in vivo* I used western blot analysis (see figure 3.4.). To this end, all FoxD3 cassettes were cloned under the control of ubiquitous pCAGGs promoters to ensure maximum expression of exogenous FoxD3 fusion proteins upon whole embryo electroporation. The western blots confirmed the Avi-tagged FoxD3 proteins were both expressed *in vivo* and specifically biotinylated by the HA-tagged BirA. Background biotin levels did not

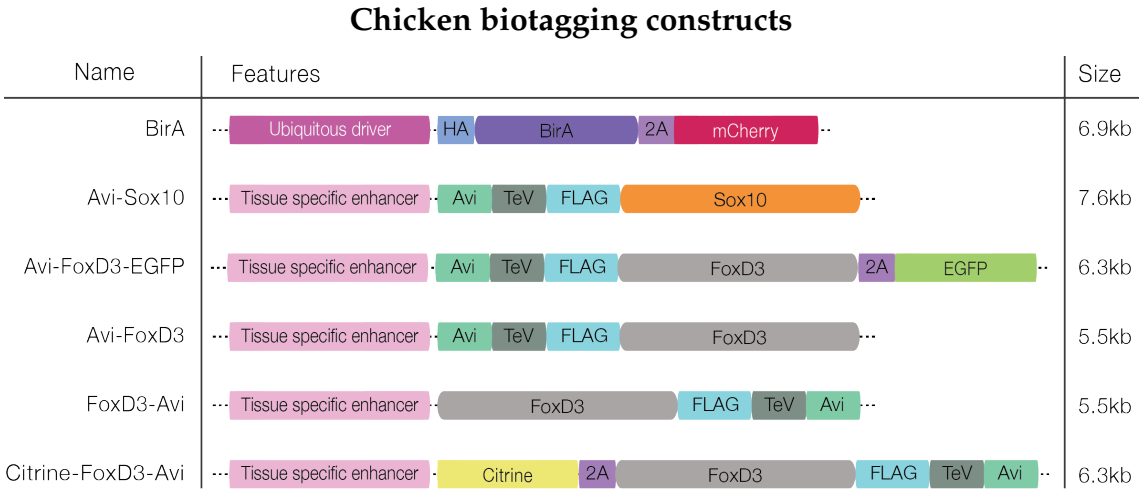


FIGURE 3.3: The biotin ligase (BirA) construct was driven by the chicken pCAGGS promoter, which results in ubiquitous expression. BirA carried a triple HA tag (HA) at the N-terminus and contained a nuclear-localisation sequence (NLS). mCherry expression provided a readout of BirA expression *in vivo*. The ribosome-skipping sequence '2A' was used to separate the mCherry transcript from the NLS-tagged BirA, thereby ensuring cytoplasmic mCherry expression. All Avi-tagged TF constructs were driven by a tissue specific enhancer acting on a ptk basal promoter and contained the Avi-tag cassette (Avi-TeV-FLAG) at either the N- or C-terminus. The Tabbaco Etch Virus (TEV) cleavage site can be used to proteolytically separate the tagged TF from the streptavidin- biotinylated Avi peptide in downstream applications (Rudra et al., 2012). The 2A ribosomal skipping peptide was used to separate EGFP and Citrine reporters in N- and C- terminally tagged FoxD3 constructs respectively.

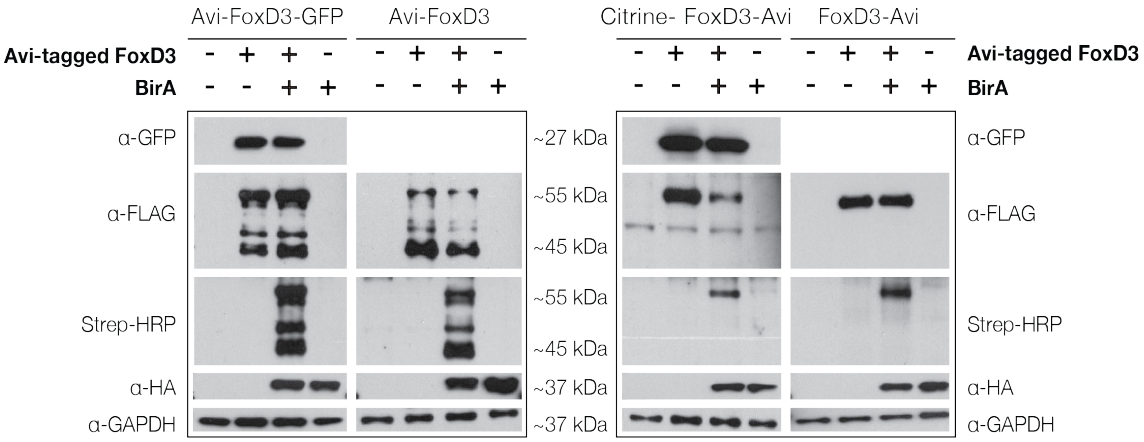


FIGURE 3.4: Detection using anti-GFP, anti-HA and anti-GAPDH showed bands at expected sizes of 27kDa, 39kDa and 37kDa respectively. Anti-HA signal was only detected in material expressing the HA-tagged BirA. Anti-FLAG signal was only detected in material expressing Avi-tagged FoxD3. Avi-tagged FoxD3 protein was expected at 55kDa, adjusted for additional tag molecular weight. Both anti-FLAG and direct Streptavidin-HRP detection showed clear differences between N- and C-terminally tagged FoxD3 protein sizes suggesting there are previously unreported variants of FoxD3 protein in chicken. All anti-FLAG and Strep-HRP bands matched in size, confirming specific biotinylation of FLAG-tagged proteins. Some background signal is seen in the Citrine-FoxD3-Avi anti-FLAG blot in all lanes.

increase due to BirA overexpression, confirming BirA was specifically biotinylating the Avi-tag.

3.2.3 Immunocytochemistry shows differences between N- and C-terminally Avi-tagged FoxD3 proteins *in vivo*

Avi-tag accessibility is essential for downstream streptavidin affinity purification. To confirm accessibility of the biotinylated Avi-tag *in vivo* and correct spatio-temporal expression of exogenous Avi/FLAG-tagged FoxD3 proteins, immunocytochemistry was used. N- and C-terminally Avi-tagged FoxD3 expression cassettes were driven by the M3 enhancer and used in combination with the ubiquitously expressed BirA

biotin ligase expressing construct. The differentially tagged constructs were bilaterally electroporated to compare the expression and *in vivo* biotinylation of the fusion proteins within the same embryo (see schematic in figure 3.1.) Citrine is a spectral variant of EGFP, and consequently both can be detected by the same anti-GFP antibodies within immunofluorescence assays.

Both BirA and Avi-tagged FoxD3 protein expression were detected in a proper spatio- temporal pattern, in addition to being localised to the expected respective cellular compartments e.g. cytoplasmic for fluorescent reporters and nuclear for Avi- and FLAG- tags. However, the anti-FLAG signal was only observed in Avi-FoxD3-EGFP expressing cells, and not in those expressing Citrine-FoxD3-Avi (see Figure 3.5.). Given detection by denaturing condition western blot, this suggests the Citrine-FoxD3-Avi has an inaccessible tag *in vivo* or that production of the Citrine-FoxD3-Avi protein was diminished due to 2A ribosome-skipping activity (Sharma et al., 2012; Trichas et al., 2008).

To test this, I compared FoxD3-Avi and Citrine-FoxD3-Avi in a bilateral electroporation experiment (see Figure 3.5.). The use of bilateral electroporation permits experimental and control conditions to be tested within the same embryo. As previously described, I did not detect FLAG signal from the Citrine-FoxD3-Avi construct but did recover anti- FLAG signal when FoxD3-Avi protein was produced without the N-terminal Citrine reporter. This indicates that the 2A component was interfering with appropriate expression of the downstream protein sequence.

I confirmed Avi-FoxD3 and Avi-FoxD3-EGFP behaved identically using bilateral electroporation (data not shown). As such, the N-terminally tagged FoxD3 constructs can be used interchangeably. Comparison of Avi-FoxD3 and FoxD3-Avi using bilateral electroporation (see Figure 3.6.) showed FLAG-tag detection on both sides of the midline. However, FoxD3-Avi showed slightly lower signal than Avi-FoxD3. The use of bilateral electroporation permits experimental and control conditions to be tested within the same embryo.

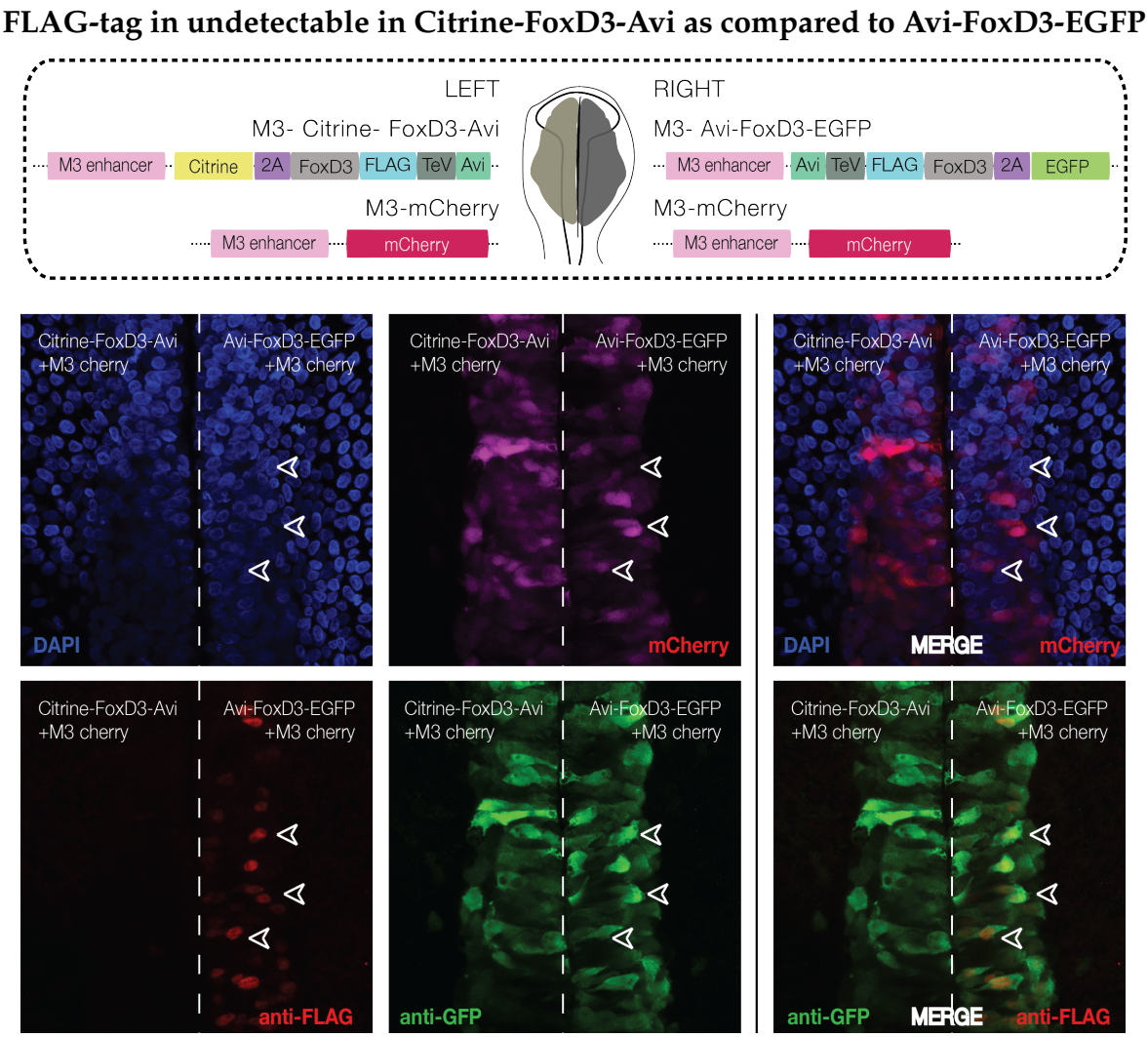


FIGURE 3.5: (A) Schematic showing the experimental design for bilateral electroporation of C- terminally tagged Citrine-FoxD3-Avi and N-terminally tagged Avi-FoxD3-EGFP on either side if an embryo. (B) High magnification views of the cranial NC of a single 8 somite-stage chicken embryo. Bilateral electroporation of Citrine-FoxD3-Avi and Avi-FoxD3-EGFP resulted in fluorescent reporter signal in the expected spatio-temporal expression pattern in the pre-migratory cranial NC. mCherry signal confirmed that the observed EGFP/Citrine expression was confined to the NC cells (arrowheads). However, while Avi-FoxD3-EGFP was detectable by the anti-FLAG antibody on the right side of the embryo (arrowheads), nuclear FLAG signal was absent on the left side of the embryo that was expressing Citrine-FoxD3-Avi. DAPI was used to stain nuclei and confirm the constructs were located appropriately.

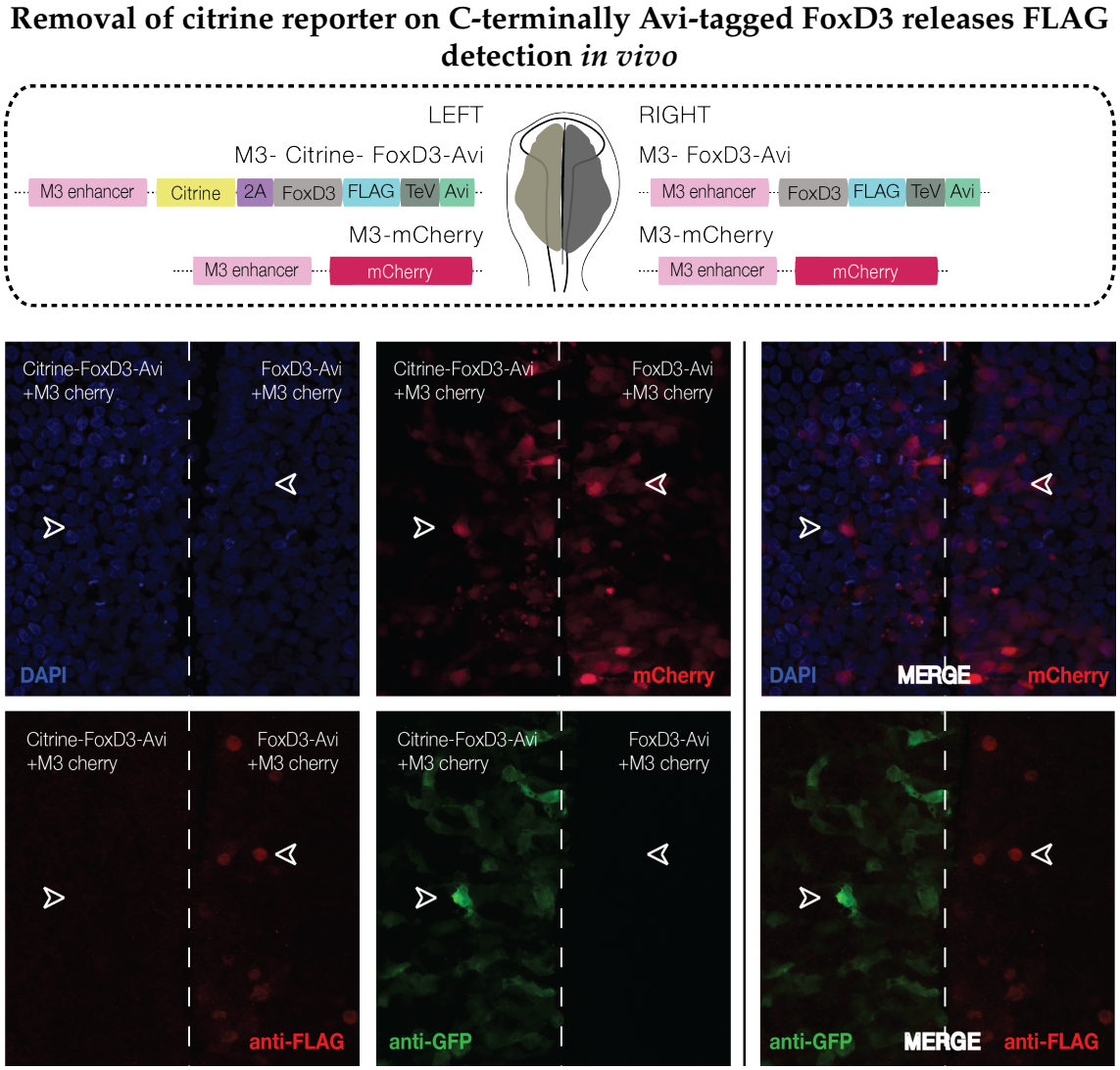


FIGURE 3.6: (A) Schematic showing the experimental design for bilateral electroporation of M3- Citrine-FoxD3-Avi and M3-FoxD3-Avi on either side if an embryo. An M3-mCherry construct was electroporated on both sides of the embryo as a positive control. (B) High magnification views of the cranial NC of a single 8 somite-stage chicken embryo. Bilateral electroporation of Citrine-FoxD3-Avi and FoxD3-Avi demonstrated interference from Citrine/2A induced ribosome-skipping. The FLAG-tag was accessible *in vivo* upon removal of the fluorescent reporter (right side arrowhead), but not accessible on side expressing the Citrine reported (left side arrowhead). mCherry expression was present on both sides indicating that the bilateral electroporation was successful. DAPI was used to stain nuclei and confirm the FoxD3-Avi was located appropriately.

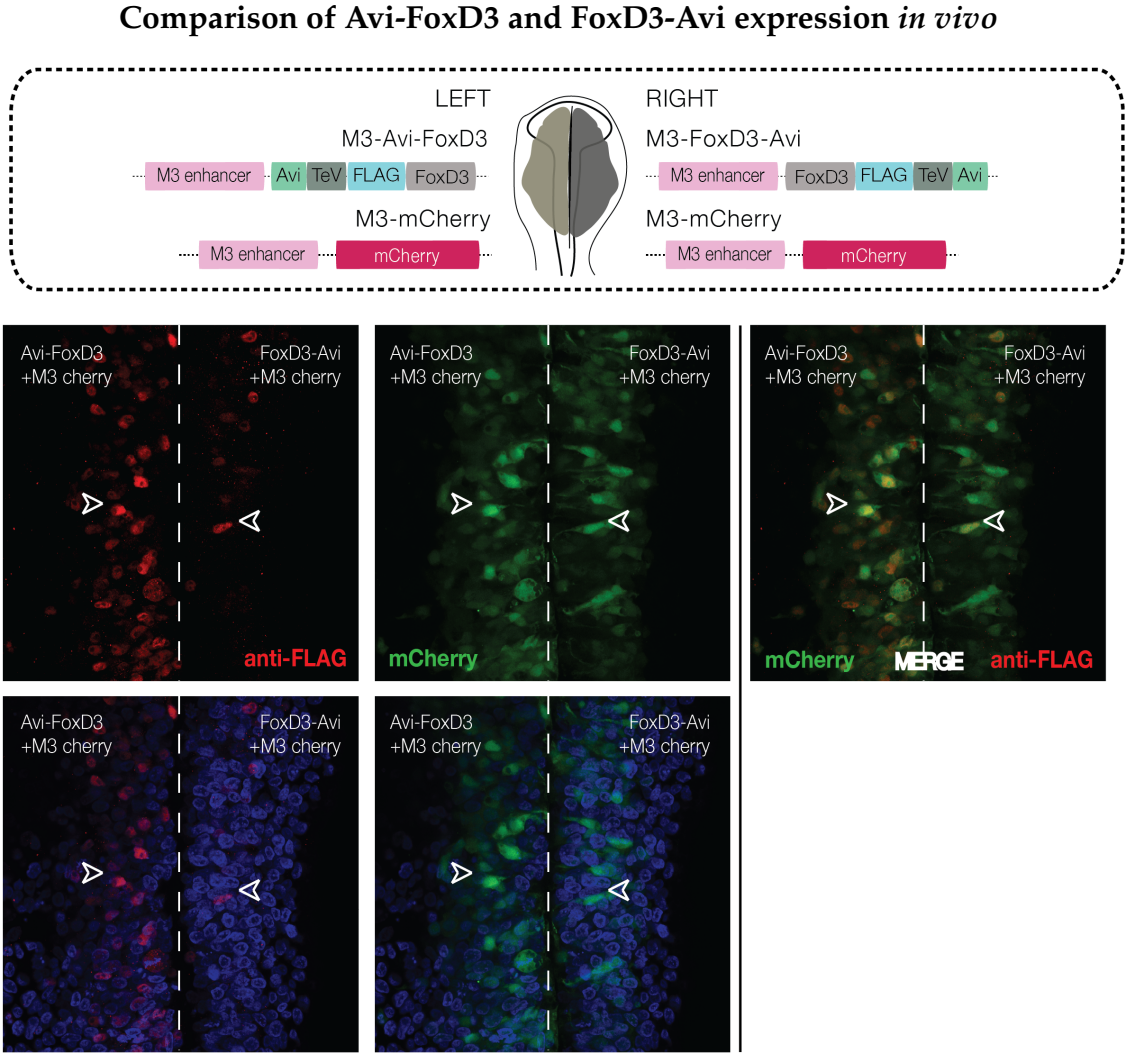


FIGURE 3.7: (A) Schematic showing the experimental design for bilateral electroporation of N- terminally tagged M3-Avi-FoxD3 and C-terminally tagged M3-FoxD3-Avi on either side if an embryo. An M3-mCherry construct was electroporated on both sides of the embryo as a positive control. (B) High magnification views of the cranial NC of a single 8 somite-stage chicken embryo. Nuclear anti-FLAG signal was observed on both sides of the midline in M3-mCherry positive NC cells (arrowheads). DAPI staining further confirmed localisation of the tagged protein to the nuclear compartment. mCherry expression was present on both sides indicating that the bilateral electroporation was successful.

3.2.4 Streptavidin accessibility of Avi- tagged FoxD3 proteins confirmed *in vivo*

I tested the accessibility of biotinylated Avi-tagged proteins to streptavidin binding *in vivo* using electroporation, subsequent streptavidin-647 immunocytochemistry and cellular resolution imaging. Given the identical expression and behaviour observed in the N terminally tagged constructs, these tests were done using the Avi-FoxD3-EGFP.

These experiments further confirmed that on a cellular level, the N-terminally Avi-tagged FoxD3 fusion proteins are only biotinylated in NC cells, and specifically by BirA (see Figure 3.8.). BirA presence does not increase overall background biotinylation in the embryo; further confirming the activity of BirA is confined to its respective Avi-tagged target. Additionally, the Avi-tags are not inappropriately biotinylated by endogenous biotin ligases.

3.2.5 Nanostring nCounter analysis showed no major perturbations of key NC genes by exogenous expression of Avi-tagged FoxD3

Nanostring nCounter technology involves the molecular barcoding of a defined code- set of target transcripts, and subsequent single molecule imaging to detect and count these barcoded transcripts (Geiss et al., 2008). Our customised code-set used included 70 important NC-TFs, signalling molecules and factors involved in development of neighbouring tissues (such as neural plate and placodes). As such, it provides a quantifiable snapshot of any perturbation of the known NC-GRN genes due to exogenous Avi-tagged FoxD3 expression.

Avi-tagged FoxD3 constructs were electroporated into HH4 chicken embryos at 1.5μg/uL final concentrations. This is within the normal range of exogenous construct concentration for testing within the chicken embryo. The effect of expressing exogenous Avi-tagged FoxD3 constructs on NC genes was assayed using Nanostring nCounter technology (see Appendix 4 for Nanostring probe set details). The

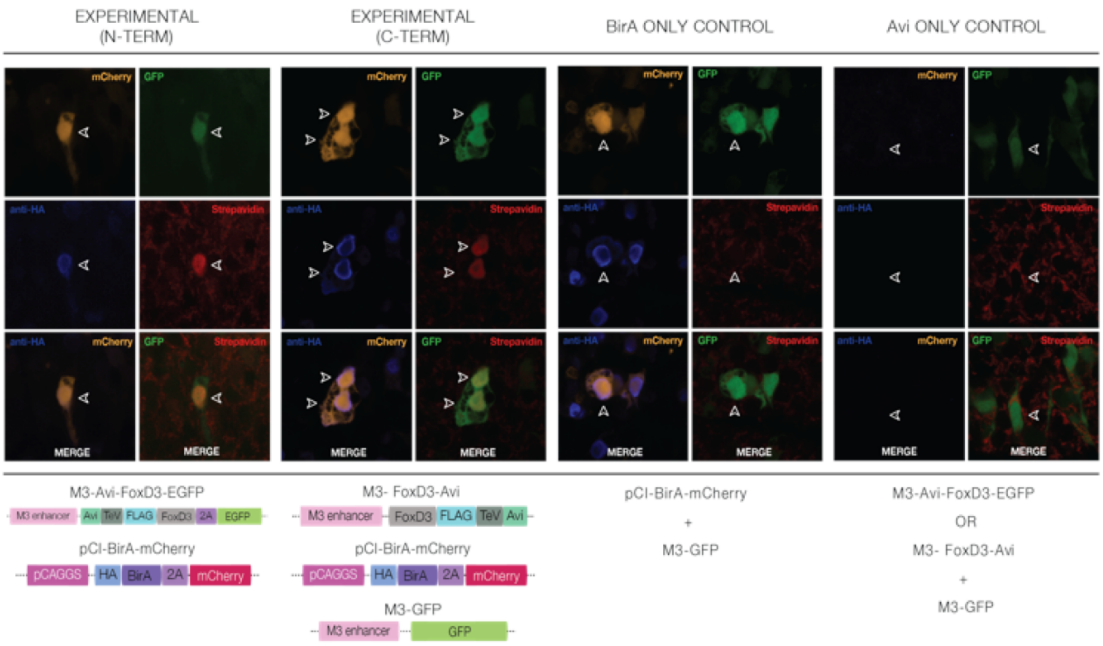


FIGURE 3.8: Cells expressing N-terminally Avi-tagged FoxD3-EGFP in addition to HA-BirA-mCherry had specific streptavidin-647 nuclear signal showing accessibility of streptavidin to the *in vivo* biotinylated protein. Cells expressing C-terminally Avi-tagged FoxD3 in addition to HA-BirA-mCherry and M3-GFP to confirm NC identity, showed nuclear streptavidin- 647 signal. The streptavidin signal was not observed in the absence of an Avi-tagged substrate or in the absence of the HA-BirA-mCherry construct.

Analysis of the effect of ectopic expression of Avi-tagged FoxD3 protein using Nanostring technology

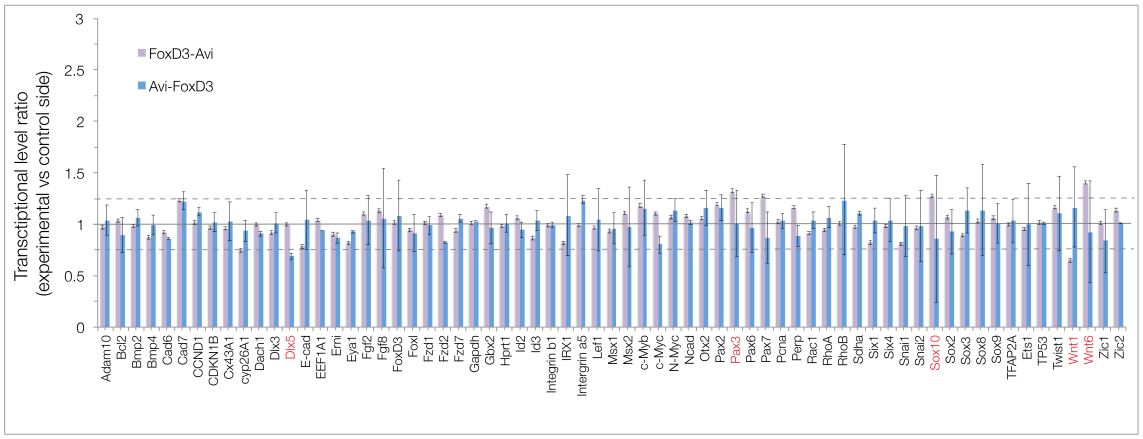


FIGURE 3.9: The effect of ectopic Avi-tagged FoxD3 expression on the wider GRN was assessed for both N- and C-terminally tagged proteins, using a NC customised code-set of 66 probes. Embryos were electroporated with BirA on both sides of the midline, and either C-terminally tagged FoxD3-Avi (grey) or N-terminally tagged Avi-FoxD3 (blue) unilaterally. Embryos were dissected at 8ss to ensure any initial changes in NC differentiation could be observed. Expression in the experimental half of the embryo was normalised against the BirA only half of the cranial region. Genes that showed a fold change of 0.25 were considered to be significantly perturbed (threshold indicated by dotted grey line). Dlx5 (green gene label) was significantly down-regulated (0.7 fold) under Avi-FoxD3 expression conditions when compared to control side but was unaffected by FoxD3-Avi expression. FoxD3-Avi had a more variable effect, with higher levels of transcription of Pax3, Pax7, Sox10 and Wnt6 and lower levels of Wnt11 in experimental material compared to control material (red gene labels). There was also more variation between the embryos assayed for the effect of FoxD3-Avi than for the Avi-FoxD3 embryos.

comparison of FoxD3-Avi and Avi-FoxD3 was repeated on two independent embryos per condition and showed little effect on the NC-GRN overall. The effect seen could reasonably be attributed to mild cellular stress or slight differences in stage between the embryos assayed (see Figure 3.9.).

3.2.6 ChIP-QPCR

I confirmed the successful affinity purification of cross-linked chromatin associated with the biotinylated Avi-tagged FoxD3 proteins using streptavidin-ChIP-QPCR. Known regions bound by FoxD3, namely the NC1 enhancer (Simões-Costa et al., 2012) and other known NC enhancers were assayed for (R.Williams, unpublished data, see Figure 5.4.).

3.3 CRISPR-Cas9 Mediated TF Knockdown in Chicken

CRISPR/Cas9 technology for targeted editing of genomic regions has made it possible to specifically modify both genomic, and more recently epigenomic landscapes, in any living organism. Classically, gene knockdown experiments in the chicken were conducted using RNA interference or morpholino-based methods (Sauka-Spengler and Bronner-Fraser, 2008). These methodologies are known to have variable levels of knockdown due to the transient nature of their mechanisms of action and are plagued by off-target effects. CRISPR-Cas9 is a programmable genome editing technique which has superseded RNAi and morpholino approaches in sensitivity, efficiency and specificity in all model systems tested in thus far.

CRISPR-Cas9 technology is a two-part system composed of a synthetic guide RNA (sgRNA) and a Cas9 endonuclease (Jinek et al., 2012; Cong et al., 2013; Mali et al., 2013). sgRNAs are composed of a short synthetic 'scaffold' sequence and a 20 nucleotide user-defined 'spacer' or 'targeting' sequence (Bolotin et al., 2005). The 'scaffold' sequence forms the hairpin docking site for the Cas9 to hold the guide RNA (Zalatan et al., 2015). The sgRNA is loaded into an endonuclease (Cas9) and ensures the targeting of the endonuclease activity to the designated target DNA sequence (Jinek et al., 2012). Cas9 induces DNA double-strand breaks (DSBs) and the cellular DNA damage complexes are consequently recruited to repair the cut site by non-homologous end joining mechanism (NHEJ) (Deltcheva et al., 2011). Indels generated at the target site result in loss of function of the encoded protein (system reviewed by Hsu et al. (2014b); Sander and Joung (2014)).

There is a single reported use of CRISPR-mediated somatic cell genome engineering in the chicken system (Veron et al., 2015). Veron *et al* reported effective knockout of Pax7 in the somite of the chicken embryo, measured using Pax7 immunofluorescence. The presented system requires the introduction of a large number of separate vectors into each cell, containing Tol2 sites flanking all cassettes to reduce plasmid dilution following cell divisions. Additionally, human U6 promoter was used to

drive sgRNA expression in chicken embryos (Veron et al., 2015).

To improve the versatility and efficiency of CRISPR-Cas9 somatic engineering in chicken, we have sought to reduce the numbers of vectors reported as required by Veron *et al* and to adapt the technology to earlier stages of chicken embryogenesis (Veron et al., 2015). Developing this technology is pertinent as loss-of-function experiments are key to confirming regulatory 'pressure points' within the GRN. We also translated recent modified Cas9 techniques for original use in chicken, to potentially elucidate the relationships between *cis*-regulatory elements and their downstream genes (Williams et al., 2018) .

3.3.1 TF knockdown in chicken via CRISPR/Cas9 mediated Non-Homologous End Joining (NHEJ)

By utilising the well-established chicken *ex-ovo* electroporation technique it is possible to introduce the CRISPR-Cas9 component vectors at HH4, prior to NC specification. During this stage, the presence of a small population of NC precursors permits the electroporation of CRISPR/Cas9 effector vectors that will have time to be expressed and to take effect. Additionally, the NC undergoes a significant population expansion at the point of EMT. Hence, the timeframe for genome engineering events to occur in the embryo for effective NC targeting is between HH4 at electroporation to 6-7 somite stage, prior to NC emigration from the NT. CRISPR/Cas9 activity is affected/enhanced by cell cycle due to the dissolution of the nuclear envelope during prophase allowing Cas9 ribonucleoprotein complex (Cas9+sgRNA) greater access to the genomic DNA.

The introduction of the genome-editing components was carried out at developmentally early stages of NCC ontogeny prior to the proliferation of NCC around EMT. Consequently, we anticipate a clonal effect from the specific CRISPR-Cas9 induced mutations in the smaller initial population of NCCs, with the entire NC population being mutated by emigration from the NT.

Non-homologous end joining (NHEJ) is a non-accurate, indel-inducing, DNA repair mechanism - the rate of which in chicken is currently unknown. As these indels are random, it remains a target of the CRISPR/Cas9 complex until an 'incorrect' indel is introduced preventing further targeting. These indels are intended to result in frameshift errors that affect protein function. By targeting Cas9 activity to either splice- donor/splice-acceptor sites, we hypothesise that the impact of any NHEJ events will be more likely to result in mis-splicing and therefore lead to gene knockout. Similarly, by targeting specific protein interaction domains, we hypothesised that the impact of NHEJ will be of greater mutagenic potential and lead to more significant knockout of the targeted TFs.

In addition to using CRISPR-Cas9 to knockout TF and look for effects on the expression of their downstream targets, observing the effect of this TF knockout on the activity of putative *cis*-regulatory elements will allow us to determine the relative importance the targeted TFs to the enhancers being characterised.

3.3.2 Constructs generated for CRISPR-Cas9 application in chicken

In addition to a standard Cas9 construct, I have generated a new chicken expression vector containing codon-optimised Cas9 followed by 2A-Citrine expression cassette (see Figure 3.10.) Optimisation to use codons preferred within the chicken system ensured efficient protein production. Next, I tested and selected the most efficient and functional U6 promoter for use in chicken embryo, to ensure the highest possible level of sgRNA delivered (see Figure 3.11.). We designed a mini-vector, separating the sgRNA expression unit from large Cas9 unit. This allowed us to deliver Cas9 independently of the sgRNAs and permitted the expression of multiple sgRNAs in a single experiment. Mini-vectors are small 3.5kb vectors, modified from pNG1-pNG4 vector backbones (Cermak et al., 2011)to contain chicken specific U6 promotors (Kudo and Sutou, 2005), and consequently are easier to deliver than larger constructs. Multiple candidate chicken U6 promoters available (RNA polymerase III)

for efficiency within the chicken system were cloned and tested. New U6-containing sgRNA vectors were designed to allow single-step Golden Gate assembly of sgRNA oligos into the vectors (Williams et al., 2018).

Key NC TF, FoxD3 that contains a single exon was used as a target to test the expression efficiency of the new U6 mini-vectors. I designed a gRNA to the forkhead domain of the protein (Figure 3.11. panel A) and cloned it into the 3.5kb sgRNA cloning mini-vectors (Figure 3.11. panel B) using BsmBI-based GoldenGate self-assembly reaction. Four different U6 promoters (U6_1, U6_2, U6_3 & U6_4), each containing the same FoxD3 sgRNA were tested for efficacy of gRNA expression (Figure 10 panel C). Individual embryos were electroporated with combinations of Cas9 plasmid and U6 vector to be tested. A minimum of 4 embryos was tested per construct. Subsequent to genomic DNA extraction from each individual embryo, I employed T7 endonuclease assays and High-Resolution Melt Analysis (HRMA) to assess the efficiency of the mismatches and indels generated by genome engineering. U6_3 was found to be the most effective. As such, all subsequent CRISPR-Cas9 based genome-engineering experiments were conducted with the U6_3 mini-vector carrying the sgRNAs of interest.

3.3.3 CRISPR-dCas9 based epigenome engineering

The CRISPR-Cas9 system can also be used to target effector proteins to specific DNA sites to engineer specific epigenetic changes at selected sites of interest. By coupling a transcriptional repressor domain such as KRAB to a catalytically inactive dead-Cas9 (dCas9) it is possible to program the down-regulation of promoters and impede enhancer activity. KRAB activity has been shown to recruit heterochromatin-forming complexes that cause histone de-acetylation and methylation (Thakore et al., 2015). Thakore *et al* 2015 reported increased H3K9me3 signal was detected at enhancers targeted by dCas9-KRAB. The effect of steric hindrance

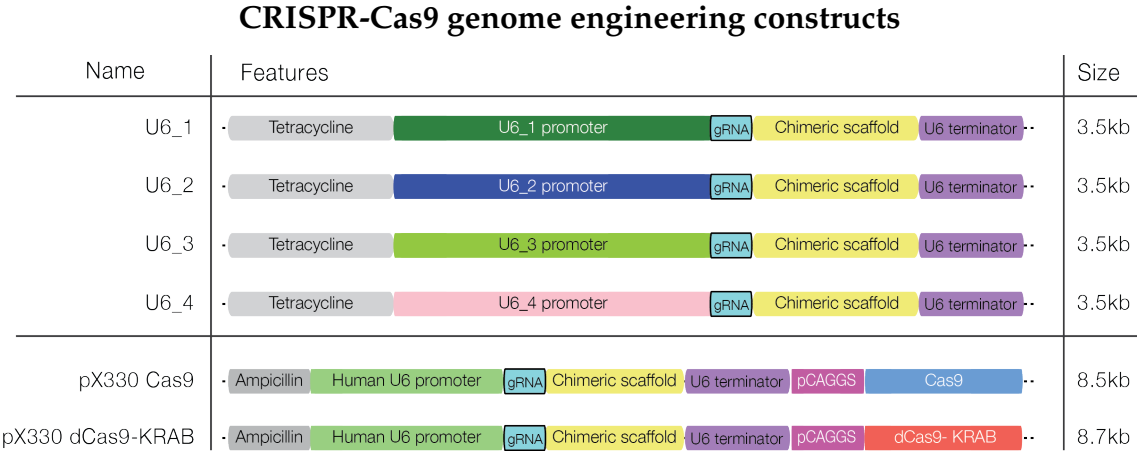


FIGURE 3.10: All chicken U6 promoter test constructs contained a gRNA GoldenGate cloning site, followed by a chimeric scaffold sequence and U6 terminator sequence. This ensured all gRNA oligos could be swiftly inserted using a GoldenGate reaction. Additionally, all U6 promoter test constructs contained Tetracycline resistance to allow for bacterial selection. The 'pX330 Cas9' construct contained a chicken codon-optimised Cas9 driven by the ubiquitous pCAGGs promoter. It also contained a Human U6 promoter to drive expression of a gRNA with a chimeric scaffold. This feature was not used in any experiments as the use of the separate chicken U6 vectors superseded its use.

as caused by the presence of the 80aa KRAB domain was noted as a possible mechanism causing the reduction of gene expression. This was confirmed by a decrease in GATA2 and FOSL1 ChIP signal at the HS2 globin enhancer assayed in dCas9-KRAB conditions as compared to dCas9-only control (Thakore et al., 2015). Additionally, they evaluated the off-target effects of dCas9-KRAB and found no additional significant sites being targeted in a global ChIP-sequencing experiment (Thakore et al., 2015). Their results showed that the local epigenomic landscape of individual enhancers could be specifically altered via dCas9-KRAB targeting (see Figure 3.12.)

3.3.4 Epigenome engineering via targeted demethylation/compaction of chromatin surrounding enhancers

The activity of *cis*-regulatory elements is facilitated by the methylation state of histones surrounding it. Custom control over epigenetic regulation of different network elements is becoming attainable with the advent of modified CRISPR/Cas9

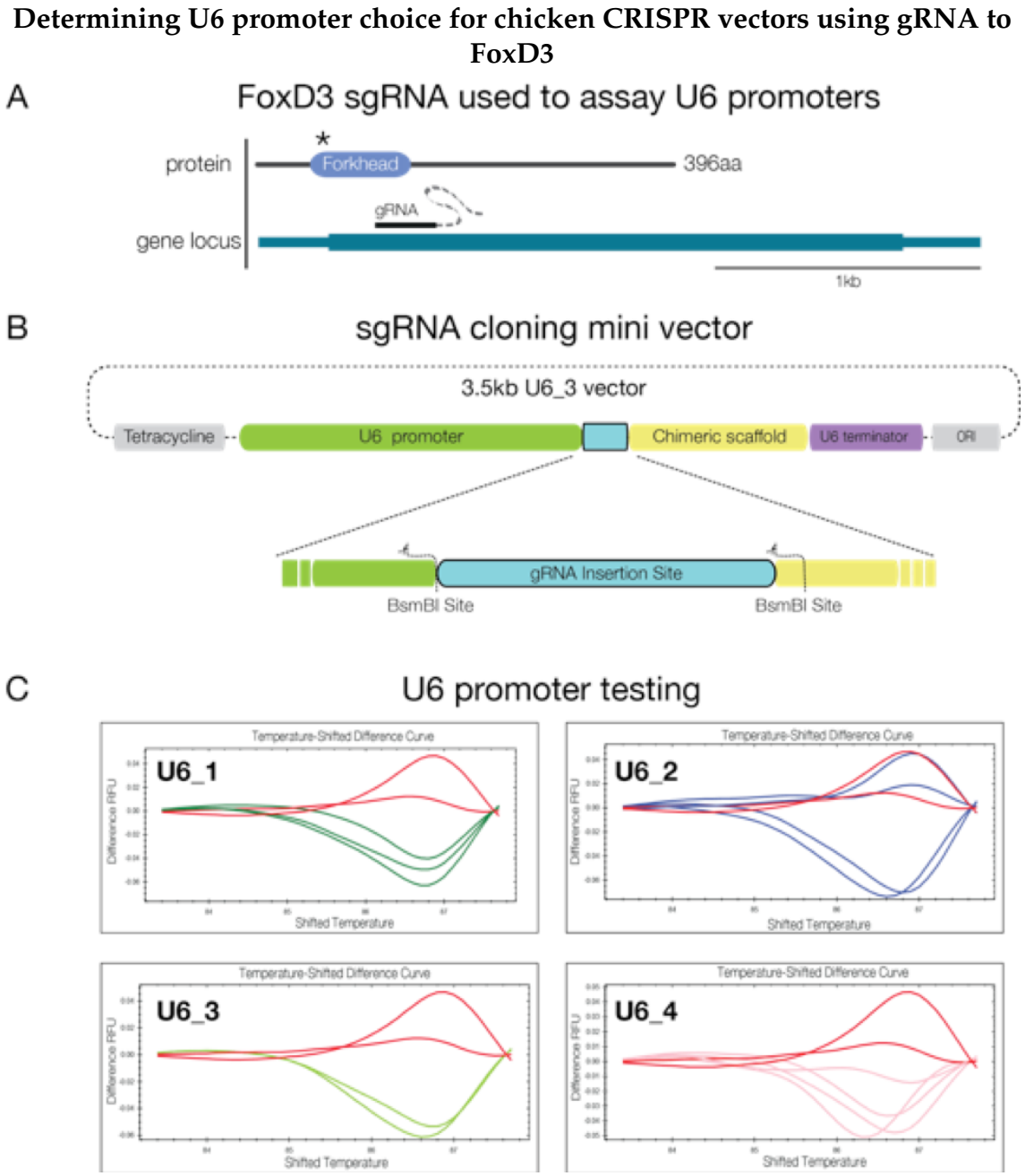


FIGURE 3.11: (A) A gRNA intended to disrupt the forkhead domain of FoxD3 was designed. (B) the gRNA was cloned into a mini cloning vector intended for use in chicken. (C) The gRNA cloned under the control of four different U6 promoters was electroporated into chicken embryos together with the pX330 Cas9 and assayed for activity using HRMA (High Resolution Melt Analysis).

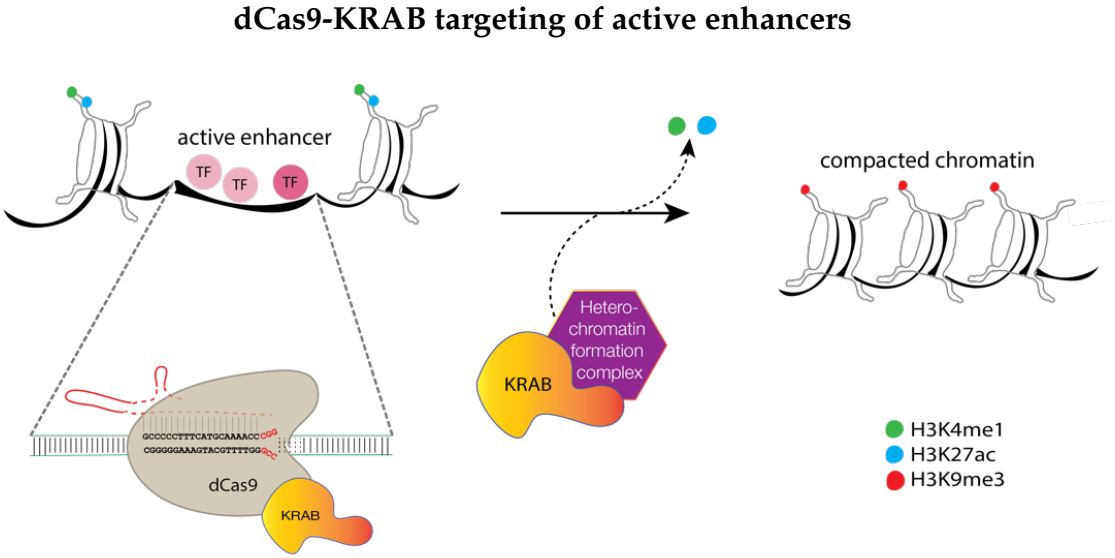


FIGURE 3.12: In combination with a gRNA to an enhancer, targeted dCas9-KRAB activity leads to the removal of histone marks (H3K4me1 and H3K27Ac) and the eventual deposition of chromatin condensing mark H3K9me3 by KRAB associated hetero-chromatin forming complexes.

technology developments. As such, dCas9-KRAB can be used to elucidate the relationship between specified regulatory elements and gene expression. I cloned a chicken codon-optimised dCas9-KRAB under the control of a ubiquitous pCAGGS promoter, in addition to a dCas9-KRAB-2A-Citrine (see Figure 3.10.). Optimisation to use codons preferred within the chicken system ensured efficient protein production. These experiments were done in collaboration with R.Williams, M. Artibani, U.Senanayake and D. Wells in the Sauka-Spengler lab (Williams et al., 2018).

3.4 Summary of Advanced Toolkit Developed to Explore Gene Regulatory Networks in the Chicken

The resolution of the NC-GRN requires the identification of *cis*-regulatory elements under the control of key NC TFs, and the establishment of the relative importance of these elements to gene expression. Increasingly, modern techniques are being applied to classical NC model systems, such as the chicken, to address previously

unresolvable questions.

Due to limited NC cell numbers at the stages of interest to us, I sought to utilise the novel zebrafish 'biotagging' technique to allow specific affinity purification of single TFs in specified spatio-temporal contexts and thereby bypass the requirement of cell sorting to acquire pure cell populations prior to ChIP-sequencing. I have successfully translated the *in vivo* biotinylation technique of 'biotagging' for use in the chicken system. Using a combination of assays the conditions and reagents required for biotin- ChIP-sequencing in the chicken have been determined.

CRISPR-Cas9 technology represents a significant opportunity to advance both the field of NC research and the applications of the chicken model system in general. As such, we sought to translate the technology into our cranial NC context. The system we developed is simple, transient and acts at an early stage of development (Williams et al., 2018). This allows experimental exploration to address questions about early NC ontogeny, in a way previously not possible. CRISPR-Cas9 induced NHEJ mutations can shed light on the activities carried out by the defined TFs, and as such contributes to the functional assessment of the NC-GRN. In addition, preliminary data suggests we are able to engineer local epigenomic changes using dCas9-KRAB at *cis*- regulatory elements, thereby assaying the relative importance of these elements on downstream gene expression.

Taken together, to facilitate the dissection of early developmental GRNs we have successfully modified two novel technologies for use in the chicken system.

Chapter 4

Identifying Putative Cranial NC Specific *Cis*-Regulatory Elements for Chromatin Remodeller Chd7 in the Chicken Embryo

4.1 Introduction

DNA polymerases, ribosomes and chromatin remodellers are key components of essential eukaryotic mechanisms such as DNA replication, RNA transcription and chromatin organisation. Mutations affecting these ubiquitous components of cellular machinery result in remarkably restricted clinical phenotypes, and frequently affect NC derived structures (Hendrich and Bickmore, 2001).

For example, mutations in DNA Polymerase E complex, involved in maintaining the stability of eukaryotic nuclear genomes, result in human 'FILS' syndrome (Pachlopnik Schmid et al., 2012). 'FILS' syndrome patients present with facial dysmorphism, immunodeficiency, livedo, and short stature. The DNA Polymerase E complex is necessary for DNA replication, cell-cycle check point control and general eukaryotic genome stability. Despite its essential function, mutations within the catalytic subunit gene of Polymerase E, POLE1, impact a surprisingly limited

range of structures (reviewed by (Pursell and Kunkel, 2008; Henninger and Pursell, 2014)).

Similarly, mutations affecting ribosomes are linked to congenital disorders including Diamond-Blackfan anaemia (DBA) syndrome, Treacher Collins syndrome (TCS), 5q- syndrome, X-linked dyskeratosis congenita, Schwachman-Diamond syndrome, and cartilage hair hypoplasia (Narla and Ebert, 2010). Interestingly, these ribosomopathies often involve craniofacial defects and/or bone marrow failure. For example, DBA, caused by mutations in ribosomal protein RPS19, is characterised by a compromised erythroid progenitor population which results in severe anaemia or bone marrow failure. Additionally, 40-62% of DBA patients present with physical abnormalities ranging from short stature to cardiac defects. Similarly, mutations in TCOF1, a constituent of peri- ribosomal ribonucleoprotein complexes, essential for ribosomal DNA transcription, result in TCS (Valdez et al., 2004). TCS is characterised by craniofacial dystosis; attributed to reduced growth of the first and second pharyngeal arch, groove, and pouch derived structures. The range of different human ribosomopathies indicates different cell-types have different sensitivities to mutation of various parts of the otherwise ubiquitous ribosomal machinery. TCS patients do not present with haematological abnormalities seen in DBA, however the craniofacial defects in TCS and DBA are highly similar despite affecting different parts of the ribosomal machinery. Taken together, the characteristic cranial dysmorphisms of ribosomopathies, and FILS syndrome patients, highlights the NC as a population of cells particularly responsive to irregularities of ubiquitous DNA polymerase and ribosomal machinery expression.

DNA methylation is another ubiquitous eukaryotic mechanism which upon disruption results in remarkably restricted aberrations. For example, CpG methylation is critical for X-inactivation, gene imprinting and embryonic development in general. These global DNA methylation patterns are established, and maintained, by the coordinated action of three ubiquitous DNA methyltransferases (DNMTs) (reviewed by (Goll and Bestor, 2005)). DNMT3b is a de novo methylase involved in

the establishment, but not maintenance, of these methylation marks. Despite ubiquitous low-level DNMT3b expression (Okano et al., 1999), hypomorphic DNMT3b mutations account for two-thirds of immunodeficiency, centromere instability, facial anomalies (ICF) syndrome cases (Maraschio et al., 1988). The facial aberrations in ICF syndrome, implicating NC hypersensitivity to DNMT3b loss-of-function, include hypertelorism, macroglossia, low- set ears, and epicanthal folds (Xu et al., 1999). Further evidence of NC sensitivity to DNMT3b activity is that ICF cells carrying DNMT3b mutations, in addition to the expected overall reduction of global methylation levels (Jeanpierre et al., 1993), additionally lack polycomb repression complex 2 (PRC2) signatures at homeobox genes critical for the development of the immune system, brain, and craniofacial region. Lentiviral knockdown of DNMT3b using a small hairpin RNA (shRNA) in H9 human embryonic stem cells alters the expression of early NC specifiers, and enhances premature NC precursor differentiation (Martins-Taylor et al., 2012). In addition to increased PAX7 and NGFR expression, DNMT3b knock-down cell lines also significantly over expressed PAX3, FOXD3, SOX10 and SNAI2, all highly significant to NC ontogeny, when compared to the mismatch shRNA expressing control cell line. This increased expression of regional identity specifying genes, and associated acceleration of differentiation, upon loss of DNMT3b is also seen in neural stem cells (Martins-Taylor et al., 2012). Taken together, the mutation of this ubiquitously expressed factor particularly reduces methylation at a limited range of sites. This suggests that DNMT3b, while a ubiquitous factor, may prefer certain genomic sites, which are of major significance only to certain tissues especially the NC.

Chromatin remodelling and DNA methylation machinery is essential for the moderation of genes underpinning cell differentiation and identity maintenance. Given their essential function, the finding that chromatin remodellers are indispensable for normal development is not particularly surprising, but it is noteworthy (Clapier and Cairns, 2009; Ho and Crabtree, 2017). However, mutations in chromatin remodeller chromodomain helicase DNA-binding protein 7 (CHD7) account

for approximately 60%-70% of human CHARGE syndrome patients (Vissers et al., 2004). CHARGE syndrome (OMIM 214800) is a sporadic, autosomal dominant disorder, characterised by a multitude of congenital aberrations in tissues and structures which can vary in both severity and presentation. 'CHARGE' is an acronym for the clinical features usually attributed to the disease: ocular coloboma, heart malformations, atresia of the choanae, retardation of growth, genital hypoplasia and ear (Blake et al., 1998; Pagon et al., 1981). Thus, despite ubiquitous expression essential chromatin remodelling function, it is striking that the effect of CHD7 mutation is limited to such few structures.

CHARGE patients also present with microcephaly, agenesis of the corpus callosum, atrophy of the cerebral cortex in addition to hypoplasia of cerebellum and olfactory bulb (Lin et al., 1990; Becker R and M, 2001; Yu et al., 2013). Chd7 is known to be indispensable for mammalian brain development (reviewed here Feng et al. (2017)) and was recently found to regulate cerebellum development via regulation of Reelin (Whittaker et al., 2017). CHD7 expression is both dynamic and spatially restricted within the brain. For example, at E12.5 and E14.5 of mouse development, Chd7 is highly expressed in medial ganglionic eminence, ventricular zone of medulla, external granule zone of cerebellum and the frontal cortex (Bosman et al., 2005). It has been noted this dynamic and restricted expression indicates tight regulation of Chd7 in the brain via unknown mechanisms (Feng et al., 2017).

The role of CHD7 in modulation of embryonic stem cell-specific gene expression is increasingly being appreciated (Schnetz et al., 2010, 2009). Interestingly, the NC is known as a stem-cell like embryonic progenitor population. It was postulated over thirty years ago that the features seen in CHARGE syndrome patients are attributable to NC dysregulation (Siebert et al., 1985), and consequently be considered a neurocristopathy (Clark, 1996). NCC proliferation within the presumptive facial primordia leads to the formation of the facial prominences which eventually fuse, forming the craniofacial structure of vertebrates (see Chapter 1.2.1 for more details). Due to this significant contribution of the NC to craniofacial structures, most

craniofacial anomalies can be attributed to the NC (reviewed by Szabo-Rogers et al. (2010)). Choanal atresia, the blockage of the nasal pathways in CHARGE patients, can be associated with holoprosencephaly and agenesis of the ethmoid bone, a NC attributed degeneration (Siebert et al., 1985). Osteoblastogenesis, a key process in the formation of the craniofacial region by NC-derivatives, has also been reported to be regulated by CHD7 (Takada et al., 2007) (reviewed by Dupin and Sommer (2012)).

The cranial placodes are specialised ectodermal thickenings, which upon differentiation, together with NCC contributions in certain placodes, give rise to components of the craniofacial region. These include the cranial sensory organs which are associated with balance, olfaction, hearing, and vision (reviewed by Streit (2004); Schlosser (2010); Steventon et al. (2014); Singh and Groves (2016)). There is extensive signalling and interplay between the NC and placodes, the surrounding mesenchyme, and the nearby ectoderm and endoderm. Consequently, many of the key features of CHARGE syndrome could be potentially linked to aberrant NC or placodal development and interplay.

For example, the olfactory placode gives rise to the migratory GnRH neurons whose dysregulation could account for the genital hypoplasia. Kallmann syndrome is a subtype of human congenital hypogonadotropic hypogonadism, in which patients additionally lack their sense of smell due to olfactory bulb abnormalities (Schwanzel-Fukuda et al., 1989). Kallmann syndrome and CHARGE syndrome are clinically distinct but share overlapping features such as hypogonadism and impaired olfaction(reviewed by Raivio (2014)). CHD7 mutations can be present in Kallmann syndrome as exemplified by patients initially diagnosed with Kallmann syndrome, but lacking mutations in expected genes (KAL1, FGFR1, PROK2, PROKR2), were found to harbour CHD7 mutations (Jongmans et al., 2008; Kim et al., 2008; Marcos et al., 2014). Heterozygous CHD7 mutant mice ($Chd7^{Gt/+}$) show dose-depended reductions in Otx2 in the olfactory placode (Layman et al., 2010) and decreased GnRH neurogenesis (Wanda S Layman and Martin, 2011). GnRH neuron specific

knockout of Otx2 result in hypogonadotropic hypogonadism in mice (Diaczok et al., 2011).

For the NC, it has been postulated that a disruption of the NC-EMT program could account for the aetiology of CHARGE syndrome (Williams, 2005). CHD7 is a chromatin remodeller and as such ubiquitous CHD7 expression by *in situ* has been reported in the chicken embryo by HH12 (Aramaki et al., 2007) and in all human foetal tissue by embryonic day 20 (Zentner et al., 2010). CHD7 chromatin remodelling activity has been shown to exert epigenetic control over NCC specification and migration both *in vivo* and *in vitro* (Bajpai et al., 2010). CHD7 knockdown in *Xenopus laevis* embryos leads to abnormal NC cell migration into the pharyngeal arches. Slug, Twist1 and Sox9, TFs critical for NC migration and specification, were severely down-regulated upon CHD7 morpholino knockdown. However, expression of NC border specifier genes Pax3, Msx1 and Zic1 was not affected (Bajpai et al., 2010). CHD7 associates with the PBAF complex at a distal Sox9 enhancer and a putative Twist1 enhancer (Bajpai et al., 2010). Previous studies have also shown that NC-specific CHD7 knockdown using a mouse Wnt1-Cre line, resulted in severe craniofacial defects similar to those seen in CHARGE syndrome (Sperry et al., 2014). Humans do not present with homozygous CHD7 mutations, and homozygosity is embryonic lethal in mice (Bosman et al., 2005). In heterozygous CHD7 gene trap mice, ectopic expression of CHD7 in NC cells alone did not rescue pharyngeal arch defects (Hurd et al., 2007). However, CHD7 overexpression in both the ectoderm and the NC did rescue genetic CHD7 knockdown; indicating that CHD7 regulates genes involved in ectoderm-to-NC signalling (Randall et al., 2009). This also indicates an important, dose-dependent role for CHD7 in the NC (Sperry et al., 2014), warranting investigation in different model systems. Together, these studies indicate that CHD7 plays a role in regulating NC genes subsequent to the neural plate border induction stage and suggest it has particular significance in EMT. Whether CHD7 regulates a single particularly sensitive NC regulator, or via direct regulation of multiple genes remains to be seen.

Neurectodermal sphere formation from *in vitro* induced human embryonic stem cells give rise to migratory cells which express NC specific genes such as AP2-a, Nestin, SOX9 and HNK1. Transplantation of these human NC like cells (hNCLCs) into NTs of chicken embryos results in their migration to the heart and craniofacial mesenchyme, differentiation into glia, melanocytes and other NC derived mesenchymal cell types. Transduction of these hNCLCs with a lentivirus encoding a shRNA against CHD7 led to a reduction of CHD7 RNA and protein levels. Subsequent to CHD7 knock-down, the migratory behaviour of these hNCLCs was significantly reduced (Bajpai et al., 2010). This further implicates the significance of CHD7 to the migration of NCCs.

Chromatin regulators investigated during human embryonic stem cell differentiation into NC progenitors versus into neurectoderm progenitors were found to be highly differentially expressed. Whilst expression levels showed >5-fold increases in these highly dynamic chromatin regulators, the two progenitor populations showed clear preferences in differentially expressed chromatin regulators (Weng et al., 2012). Chromatin regulators KDM4A, CHD7 and DNMT3A have been shown to have high levels of expression at early stages of NC ontogeny prior to EMT (Bajpai et al., 2010; Hu et al., 2012; Strobl-Mazzulla, Pablo Hernan et al., 2010). This is reflected in the similarity between ICF, hypomorphic DNMT3b mutations leading to characteristic facial aberrations, and CHARGE syndromes respectively (Maraschio et al., 1988; Pauli et al., 2017; Vissers et al., 2004). This type of differential expression indicates that these chromatin modifiers may be under tight transcriptional control; an end usually achieved through the activity and plenitude of enhancer elements. It is also possible that a bystander effect from neighbouring genes could also account for this increased expression. However, the apparent sensitivity of the NC to mutations affecting the expression of these chromatin regulators would suggest this is not a bystander effect of neighbouring gene expression (Hendrich and Bickmore, 2001). It is noteworthy that the mutation of two different chromatin regulators account for a significant proportion of human syndromes characterised by NC dysregulation.

Increasing clinical evidence suggests that there are human diseases with causative mutations in *cis*-regulatory regions (Epstein, 2009). For example, mutations within a regulatory region of the RET gene has been reported as causative of Hirschprung's disease (Emison et al., 2005). Hirschprung's disease is attributed to mutations within the RET gene and is known to occur due to aberrant trunk NC development (reviewed by (Arshad et al., 2012)). Alterations in TF binding sites within enhancers have been shown to have clinical relevance in the cranial NC also. For example, non-syndromic cleft lip is strongly associated with SNPs in the IRF6, interferon regulatory factor 6, locus. Furthermore, a cleft lip/palate causative SNP was identified in a novel IRF6 enhancer, which interfered with an AP2a binding site (Rahimov et al., 2008). Mutations affecting an MSX1 binding site within a SOX9 enhancer were found to be highly implicated in the Pierre Robin sequence, another human syndrome with NC aberrations similar to those found in CHARGE syndrome (Benko et al., 2009).

CHD7 expression is initially ubiquitous in human (Sanlaville, D et al., 2006), mouse (Bosman et al., 2005; Lalani et al., 2006) and throughout the neural epithelium during neurulation of chicken embryos (Aramaki et al., 2007). Human foetal tissues show ubiquitous CHD7 expression by 20 days of development, restricting to the central nervous system and NC containing mesenchyme of the pharyngeal arches from days 26-33. At day 33, CHD7 expression in human foetal tissue is also observed in the otic vesicle but is noted as having stronger *in situ* signal in the spinal cord and NC-derived dorsal root ganglia (Sanlaville, D et al., 2006). While Sanlaville *et al* 2006 reported an absence of CHD7 expression in human foetal heart tissue, Aramaki *et al* 2007 noted that the stage 20 chicken Chd7 expression in the NC-populated branchial arches could account for the heart defects tissues, for example the ear, eye and olfactory system (Aramaki et al., 2007). Additionally, the dynamic restriction of CHD7 from ubiquitous to tissue-specific expression is also temporally regulated, strongly indicating that *cis*-regulatory elements like enhancers to CHD7 should exist in CHARGE-relevant tissues (Zentner et al., 2010). The disruption of

elements providing spatio-temporal regulation of CHD7 could be of clinically relevant in CHARGE syndrome.

As chromatin regulators are ubiquitously expressed (Nestorov et al., 2015), it has been assumed that cells utilise the same 'housekeeping' machinery to constitutively transcribe chromatin regulators and remodellers. In contrast, tissue-specificity of developmental regulators, or terminal differentiation factors, is thought to be achieved by cooperative activity of tissue-specific TFs acting upon tissue-specific enhancer elements, open and accessible only in specific cell types. Therefore, the possibility of tissue-specific enhancers for chromatin regulators such as CHD7 is highly intriguing.

For the remaining 25-40% of CHARGE syndrome patients with no mutation in the CHD7 coding region ((Vissers et al., 2004; Blake and Prasad, 2006; Lalani et al., 2006)), it has been hypothesised that mutations in either the 5' or 3' untranslated regions of CHD7 or in thus far unknown critical distant *cis*-regulatory elements may be responsible for the phenotype. Spatio-temporal regulation of CHD7 by tissue-specific enhancers could account for the restriction of initial ubiquitous embryonic expression of the gene only to structures affected in CHARGE syndrome - e.g. the eye, ear and olfactory organs (Zentner et al., 2010). A recent review of CHD7 in the brain noted that tight regulation of CHD7, via unknown mechanisms, must give rise to the particular spatio-temporal expression patterns of CHD7 within the brain (Feng et al., 2017). Thus, identifying a complement of CHD7 enhancers could be instrumental in identifying the possible mutations in the unaccounted 40% of CHARGE patients with no CHD7 mutations in the gene exons. Given the NC aberrations in CHARGE syndrome, and the indications that CHD7 expression could be tissue-specifically restricted to certain tissues through development, we hypothesise the existence of NC-specific enhancers for this factor. The identification of NC-specific Chd7 enhancers would complement the resolution of the wider NC-GRN, in addition to potentially providing sites to screen for causative mutations in non-CHD7 mutated CHARGE patients. As it has been suggested that the role of CHD7

in the NC might be at its most significant prior to final NC differentiation (Weng et al., 2012), I intend to identify active enhancers during NC- EMT.

In this chapter, I investigate the possible tissue-specific regulation of chromatin remodeller Chd7 during chicken embryonic development. It is surprising that the loss of ubiquitously expressed chromatin regulator CHD7 has the limited and ostensibly tissue-specific effect seen in human CHARGE-syndrome. NC dysregulation is particularly implicated in the CHARGE syndrome phenotype. As such, we hypothesised that NC regulation of CHD7 would require NC-specific enhancers controlled by tissue-specific factors. Here, I identify and test putative enhancers for Chd7 in the NC, using traditional chicken fluorescent reporter assays. Subsequent to FACS isolation of cranial NCCs, ATAC-sequencing was used to highlight accessible sites. These accessible sites are enriched for enhancer elements and correlate with cell-type specific histone modifications and cell-type specific gene expression (Xi et al., 2007). Enhancer and promoter interactions occur within the framework of topologically associated domains (TADs). TADs are stable units of genomic architecture, thought to partition genomes into large regulatory units, which are maintained by CTCF binding (Dixon et al., 2015). CTCF binding sites are found at the edges of TADs, and their disruption leads to the subsequent disruption of the TADs they otherwise maintain. Disruptions to TADs can lead to aberrant enhancer-promoter interactions, leading to misexpression of said genes, and can cause malformation syndromes (Lupiáñez et al., 2015). All the enhancers tested lie within the TAD surrounding CHD7 (unpublished, R.Williams), and consequently if active, will most likely regulate CHD7.

4.2 Results

4.2.1 Onset of Chd7 expression overlaps with key NC TFs FoxD3 and Sox10 in the chicken embryo

To establish the expression pattern of Chd7 in the chicken, and confirm expression in the NC, I used *in situ* hybridisation to compare the expression profile of Chd7 to that of key NC TFs FoxD3 and Sox10 (see Figure 4.1.). FoxD3 marks bona-fide specified NC, and Sox10 marks the post-EMT and migrating NC (Gammill, 2002; Sasai et al., 2001; Simões-Costa et al., 2012). Using DIG-labelled RNA probes, embryos between HH stage 8 and 11 were probed for Chd7, FoxD3 and Sox10 transcripts (for probe details see Appendix 5).

At HH8, Chd7 expression appears localised to the neural folds. While maximum signal intensity remained in the NT, by HH11 expression was observed throughout the embryo. The Chd7 expression pattern observed was consistent with that previously published in chicken (Aramaki et al., 2007). The observed expression pattern is congruent with previous reports of low-level ubiquitous CHD7 expression in other organisms (Vissers et al., 2004; Schulz et al., 2014). Furthermore, the lack of signal in the sense probe condition supports the observation of the antisense probe seen in an apparently relatively ubiquitous way. FoxD3 and Sox10 expression patterns were as previously published (Khudyakov and Bronner-Fraser, 2009). Chd7 expression overlapped with that of FoxD3 at the onset of FoxD3 expression at HH stage 8 in the neural folds (indicated by arrows in Figure 4.1.). Chd7 expression also overlapped with that of Sox10 at the onset of Sox10 expression at HH stage 8 in the neural folds (indicated by arrows in Figure 4.1.). At HH11, Chd7 expression overlaps with, and extends beyond, the distinct 'cobra-head' Sox10 expression pattern.

Together, these *in situ* hybridisations confirmed Chd7 expression in the neural folds at HH8, a spatio-temporal region known to harbour the newly specified NC. Interestingly, the expression of Chd7 commences in this NC containing zone, and

the NT consistently shows strongest signal compared to other tissues.

4.2.2 Comparative chromatin accessibility profiling indicates multiple putative *cis*-regulatory regions within the epigenomic landscape surrounding Chd7 in cranial NCCs.

To identify putative cranial NC specific enhancers for Chd7, comparative epigenetic profiling of NCCs to non-NCCs was conducted. As enhancers encode both spatial and temporal information, regions potentially being employed in a tissue-specific manner were highlighted by comparing the epigenetic profiles of NCCs and non-NC cells. The temporal relevance these regions was demonstrated by comparing epigenetic profiles at different time points in NCC ontogeny, namely 10-12ss and HH4 epiblast stages.

Chicken embryos were electroporated with the cranial NC specific M3 enhancer known to drive GFP reporter activity in pre-EMT and migratory NCCs (Simões-Costa et al., 2012). Upon reaching 10-12ss, embryos were dissected and dissociated into a single-cell suspension and the fluorophore expressing cranial NCCs isolated from the environing cranial mesenchyme non-NC cells via Fluorescence Activated Cell-sorting (FACS). Subsequent to separating the fluorescent cranial NCCs from the non-fluorescent non-NCCs, ATAC-seq conducted upon them. ATAC-seq is a chromatin accessibility assay which highlights accessible areas within the genome, presumed to contain sites such as enhancers and other *cis*-regulatory elements (Buenrostro et al., 2013). RNA-sequencing was also conducted upon FACS isolated cranial NCCs at 10-12ss and further confirmed expression of Chd7 in the NC (see RNA-seq track in Figure 4.2.) (R.Williams, unpublished).

To identify putative regulatory regions, ATAC-seq profiles for 10-12ss FACS isolated NCCs were compared to control non-reporter expressing cranial mesenchyme cells. Regions with low accessibility peaks in control cells compared to the NC population were deemed putative regulatory regions and taken for further investigation

Expression of FoxD3, Sox10 and Chd7 in the cranial region of chicken embryos at HH stage 8 to 11

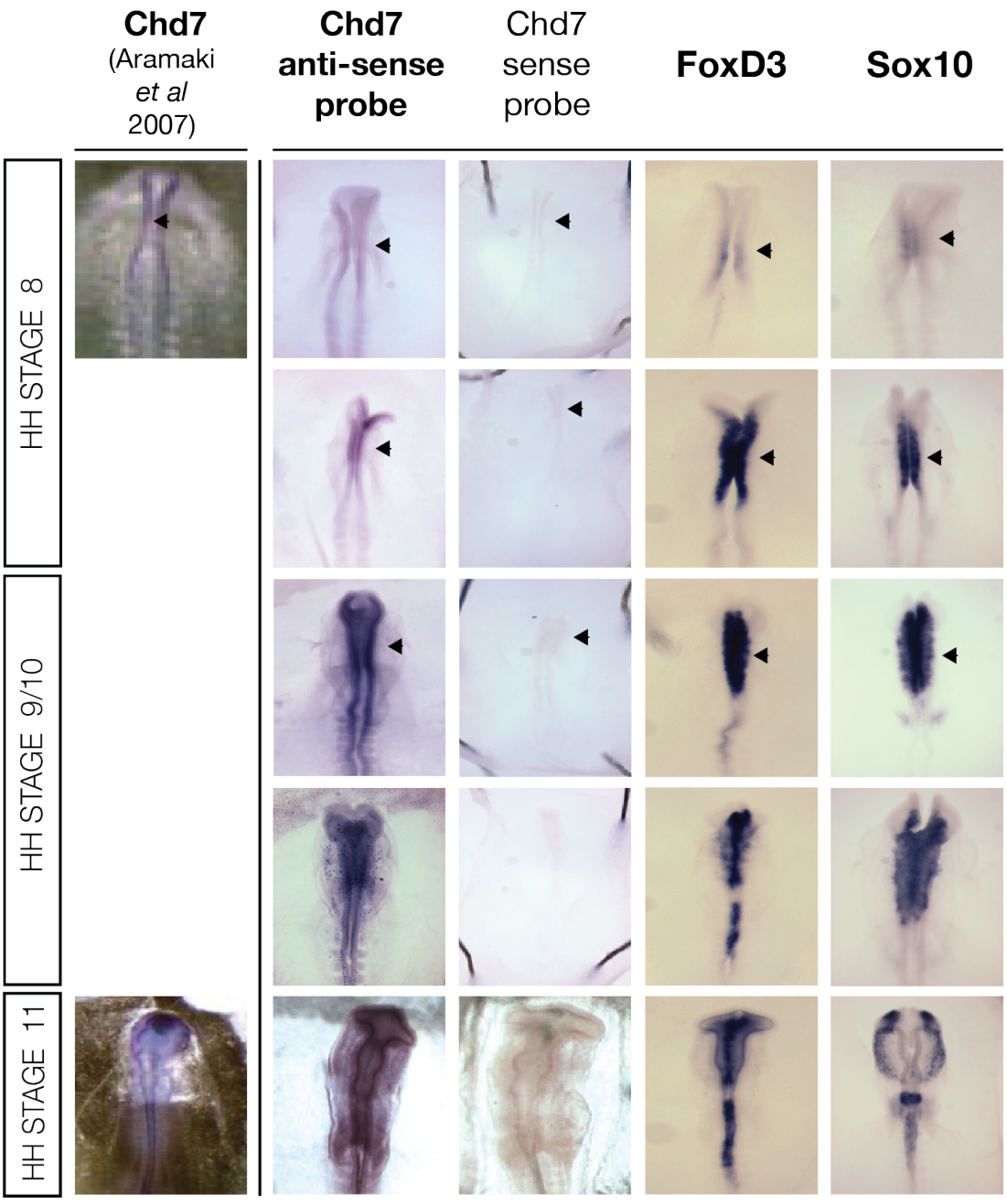


FIGURE 4.1: Gene expression patterns were analysed by whole mount *in situ* hybridisation using anti-sense digoxigenin-labelled RNA probes for FoxD3, Sox10 and Chd7. Early during HH stage 8, both FoxD3 and Sox10 expression was confined to the cranial neural folds (arrowheads). Later in HH stage 8, FoxD3 and Sox10 expression intensified in the neural folds and Chd7 expression was visible in a similar region. At HH stage 9/10, FoxD3 and Sox10 expression were confined to the pre-migratory and migrating NC cells (arrowheads). In contrast, Chd7 expression spread throughout the entire chicken embryo. At HH stage 11, FoxD3 expression was confined to the neural tube. Sox10.was expressed in the migrating NC as well at the neural tube. Chd7 was clearly expressed throughout the embryo. The observed Chd7 signal was in keeping with that in the only report of Chd7 expression in chicken (Aramaki et al., 2007). A sense probe for Chd7 was used to confirm the specificity of the signal observed with the corresponding anti-sense probe

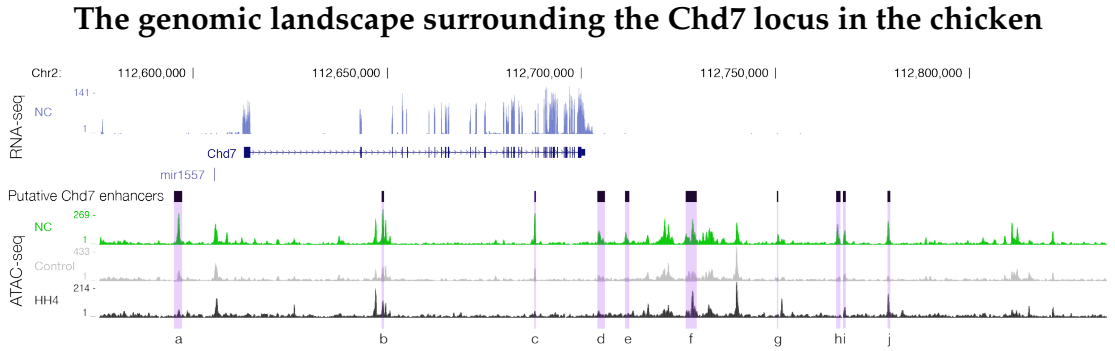


FIGURE 4.2: RNA-seq data (blue track) from NC1-GFP sorted NC cells around EMT confirmed that Chd7 is expressed in the NC ATAC-seq profiles show the regions of open chromatin at the Chd7 locus in 8-10ss NC1-GFP sorted NC cells (green), GFP-negative control cells at the same stage (grey) and cells from HH4 stage embryos (black). Putative cranial NC-specific enhancers were selected by comparing the ATAC-seq profiles. Regions highlighted in purple were selected as putative enhancers and tested in reporter expression assays by electroporation of chicken embryos.

(see Figure 4.2.). Only peaks within the Chd7 TAD were considered within the scope of this project. This was due to the rule-of-thumb that enhancers for a gene are usually found in the area encompassing the gene in question, between the neighbouring upstream and downstream genes (Furlong and Levine, 2018). Further confirmation of the temporal aspect of the putative enhancers was the comparison to the epigenetic profile of HH4 epiblast stage whole embryos. The HH4 ATACseq profiles were generated from dissociated whole epiblasts (R.Williams unpublished).

Ten putative cranial NC Chd7 *cis*-regulatory regions were identified using this method (for details of putative enhancers see Figure 4.3.).

4.2.3 Seven enhancers identified in a screening of putative cranial NC specific Chd7 *cis*-regulatory regions

The comparison of chromatin accessibility at different stages of NC ontogeny highlighted ten potential NC specific Chd7 *cis*-regulatory regions (see Figure 4.4.). To establish if any of the putative Chd7 *cis*-regulatory regions contained enhancer elements, the putative elements were cloned into fluorescent reporter constructs containing a minimal promoter upstream of a fluorescent protein (see Appendix 1 for

Table of putative Chd7 NC enhancers

Name	Size (bp)	Distance to Chd7 TSS	Chromosome location
Chd7_ a	2104	-17.3 Kb	chr2:112,594,568-112,596,672
Chd7_ b	696	+35.9 Kb	chr2:112,648,524-112,649,219
Chd7_ c	380	+75.0 Kb	chr2:112,687,882-112,688,261
Chd7_ d	1616	+92.0 Kb	chr2:112,704,220-112,705,835
Chd7_ e	707	+98.5 Kb	chr2:112,711,257-112,711,963
Chd7_ f	2949	+115.3 Kb	chr2:112,726,796-112,729,745
Chd7_ g	414	+137.7 Kb	chr2:112,750,489-112,750,902
Chd7_ h	993	+153.3 Kb	chr2:112,765,767-112,766,759
Chd7_ i	806	+154.9 Kb	chr2:112,767,559-112,768,364
Chd7_ j	518	+166.0 Kb	chr2:112,778,855-112,779,372

FIGURE 4.3: Ten putative enhancers were cloned into basal ptk promoter containing reporter expression constructs. The largest region cloned was Chd7_f (2.9kb) and the smallest Chd7_c (380bp). The majority of the cloned regions were located downstream of Chd7; reflecting the Capture-C identified topologically-associated domain (R.Williams, unpublished).

further details). These constructs were electroporated into chicken embryos (see Figure 4.3.) and observed for enhancer-driven reporter activity from 5ss onwards (see Figure 4.4.). To confirm NC-specific activity of the putative *cis*-regulatory regions being tested, all reporter constructs were co-electroporated with a NC-specific reporter construct and assayed for overlapping fluorescence. This NC-specific reporter construct highlights both recently-specified and migrating cranial NCCs as it contains the M3 enhancer of FoxD3 upstream of a minimal promoter and fluorescent reporter.

Of the ten *cis*-regulatory regions tested, seven were found to have enhancer activity as reporter signal was detected. Of the seven active enhancers identified, one appeared ubiquitous, whilst the others had distinct spatial patterns.

The Chd7_a enhancer was strongly active in the cranial NC at HH 10 onwards and was consistently driven in M3-expressing NC cells (Figure 4-4 panel 1, n=29/29). Enhancer Chd7_c showed strong activity in the trunk NC at HH stage 8 (white arrows Figure 4.4.), which later spread into the midbrain and a small number of cranial NC cells (white arrows Figure 4.4. panel X, n=11/11). Chd7_d also showed enhancing activity in the cranial region, and other neighbouring cranial tissues (Figure 4.4. panel X, n=10/10). Chd7_f showed specific activity in M3-expressing cranial NC cells, with the onset of expression at HH 8-9 (Figure 4.4., n=12/12). Chd7_h had activity in the cranial NC, in addition to a small number of other M3-negative cells (Figure 4.4. n=12/12). Chd7_i had some activity in the cranial and vagal regions (Figure 4.4. n=5/5). Chd7_j showed strong expression throughout the entire embryo at all stages observed (Figure 4-4 panel 4, n=12/12). Putative enhancers Chd7_b (n=5/5), Chd7_e (n=9/9) and Chd7_g (see Figure 4.4.) were negative for enhancing activity in the chicken reporter assays conducted around the stage of NC-EMT.

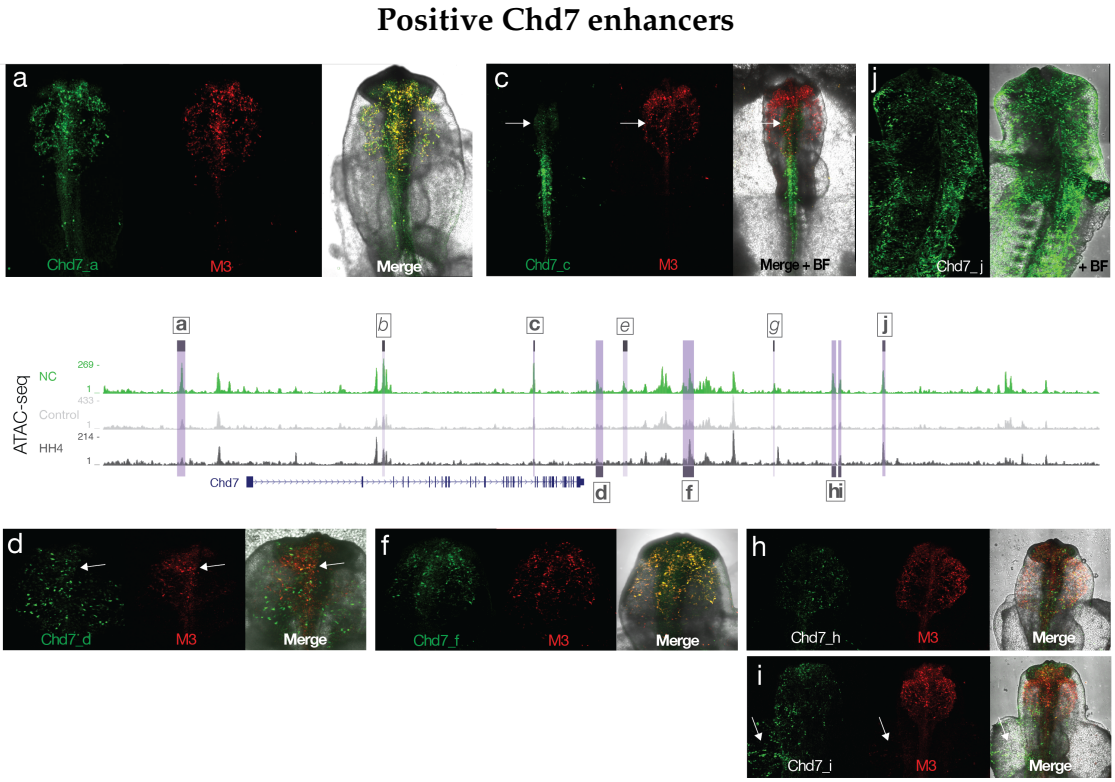


FIGURE 4.4: Purple bands highlight the regions of open chromatin at the chicken Chd7 locus, revealed by comparison of ATAC-seq profiles for 8-10ss NC1-GFP sorted NC cells (green track), GFP-negative control cells at the same stage (grey track) and cells from HH4 stage embryos (black track). Ten putative enhancers (black boxes and corresponding shaded regions) were identified around the Chd7 locus and tested in reporter expression electroporation assays along with an M3 construct to label the NC cells. Seven of these enhancers showed GFP (enhancers c, j, d, h, i) or Cerulean (enhancers a and f) reporter activity. All images show the cranial region of HH8-11 embryos. Six enhancers showed NC tissue-specificity (panels corresponding to enhancers a, c, d, f, h, and i) and Chd7_j drove ubiquitous reporter expression (panel j). Out of the putative enhancers Chd7_c (panel c) is the only intronic enhancer, located upstream of exon 24. White arrowheads indicate regions of putative enhancer reporter and M3 expression overlap. In GFP enhancer reporter conditions the M3-NCC reporter construct carried mCherry activity, in cerulean enhancer reporter conditions the M3- NCC report construct carried GFP activity.

4.3 Discussion

In light of the hypothesis that some causative CHARGE syndrome mutations could be located in CHD7 *cis*-regulatory regions, I profiled Chd7 expression in the 6-12ss chicken embryo and identified a number of putative chicken NC specific Chd7 *cis*-regulatory sequences.

In situ hybridisation was used to determine the expression of Chd7 in chicken as this was previously unknown in finer detail across HH8-11. As per the published pattern (see published Chd7 column in Figure 4-1), the Chd7 expression pattern observed changed from initial restriction to the neural folds at HH8 to apparent ubiquitous low-level expression by HH11. This ubiquitous but low-level signal was corroborated by the sense probe embryos being entirely devoid of signal. The initial limited Chd7 expression indicates that tissue-specific *cis*-regulation of Chd7 in chicken might occur, as this region is also the site of NC specification. Additionally, at this early stage the Chd7 expression pattern overlapped significantly with the expression patterns of known key NC specifying TFs- FoxD3 and Sox10. RNA- sequencing of NC1-FoxD3 FACS isolated cells showed high Chd7 expression (R.Williams, unpublished). Further proof of Chd7, FoxD3 and Sox10 co-expression could be demonstrated by coronal sections showing overlapping enhancer reporter activity, *in situ* or RT-QPCR mRNA detection, or immunohistochemistry. The distinct localisation of Chd7 expression onset to the neural folds is interesting and should be noted as being potentially significant as CHD7 expression remains distinctly higher in this localisation throughout the timeframe assessed and is known to be of significance to neural derivatives and development.

Using comparative ATAC-seq profiling, to identify regions of dynamic accessibility, ten putative Chd7 regulatory regions were selected and tested for enhancing activity. This dynamism was established using ATAC-profiles from two stages of NC ontogeny between HH4 and HH10-12, as compared to a control population on

non-NC cells. These putative enhancers were cloned into fluorescent reporter constructs and screened in vivo for enhancing activity by electroporation into chicken embryos. The known NC- specific FoxD3 enhancer M3 was used to highlight the NC, thereby establishing if any putative enhancer activity was seen in the NC. Capture-C (Davies et al., 2016; Hay et al., 2016; Hughes et al., 2014) from the Chd7 promoter has been done within our lab (R. Williams, unpublished). All ten putative regulatory sites selected fall within the topologically associated domain of the Chd7 promoter. This strengthens the presumption that these enhancers regulate Chd7 as they are enclosed in a domain. However, capture from each respective enhancer would be further proof of Chd7 regulation.

The trunk NC like reporter expression driven by Chd7_c, could be confirmed by co-electroporation with the known 'NC2' trunk FoxD3 enhancer (Simões-Costa et al., 2012). Novel enhancer Chd7_d had expression beyond that of only the cranial NC, but this region contains a double peak. As such, Chd7_d could be refined further, and potentially restricted to the NC- specific component. The strong and NC-specific enhancer Chd7_f also contains a double ATAC-seq peak profile in the 2.9kb region. These could be tested separately to establish the relative contributions from each portion to the observed reporter expression. The Chd7_i enhancer appears to have a vagal NC contribution. The trunk and vagal NC-specific activity could be established via co-electroporation of other known trunk and vagal NC enhancers where possible (Chan et al., 2005; Simões-Costa et al., 2012). Further characterisation with known vagal markers, e.g. RET, or observation of migration patterns, e.g. gut ganglia formation etc, could be used to confirm the fluorescent cells are vagal NCCs. The Chd7_j element had ATAC-signal in both NCCs and non-NCCs. The apparent ubiquitous activity indicates that this enhancer can be utilised in a wide range of tissues and is not limited to the NC.

Further characterisation of all these novel enhancers would be aided by sectioning to assay which layer, and cells, expressed them. The expression patterns of each of these enhancers should be observed at later stages to further characterise

the tissue-specificity and then determine which NC derivatives in particular use the enhancers. However, these results are an indication that Chd7 is regulated by a number of different combinations of factors in different cranial cells. It would be interesting to investigate the possibility that differential Chd7 expression is indicative of different cranial NC derivatives.

Three regions out of the ten selected were found to be negative for enhancing activity between HH8 and 10 in the embryos assayed (Chd7_b, e and g). It is possible that these regions are *cis*-regulatory elements for Chd7 at different stages of development. Alternatively, the accessible sites unable to drive fluorescent reporter activity together with the minimal promoter in the reporter construct, could be other structurally relevant elements. For example these regions could contain CTCF sites known to be essential for correct chromatin looping and TAD formation and maintenance (Guo et al., 2015; Guo and Qiang, 2015).

In summary, I have successfully identified two strong cranial NC specific enhancers using the combination of ATAC-seq profiles and chicken reporter assays (ATAC-seq profiles generated by R. Williams, unpublished). Five additional novel Chd7 enhancers have also been identified, including one putative trunk NC enhancer, a putative vagal NC enhancer and a ubiquitous enhancer. Most excitingly, this data shows that tissue- specific *cis*-regulatory elements for chromatin remodeller Chd7 exist in the chicken, and importantly that these findings could be of clinical relevance due to the similarities in chromatin landscape of humans.

Future work includes looking for the presence of these regions within the human genome. Whilst sequence conservation might not be identified, it is possible that there might be syntenic, or functionally, conserved regions in similar locations of the human Chd7 locus. If any of these novel enhancers are conserved in humans, they would represent sites potentially harbouring CHARGE syndrome causative mutations in patients not carrying Chd7 exonic mutations. Specifically, one might anticipate enhancer only mutations would present with a milder CHARGE syndrome phenotype as it would only impact CHD7 expression in a sub-population

of cells employing said enhancer. More specifically, enhancer only mutations might be predicted to present as a more restricted phenotype, perhaps lacking some of the congenital malformations usually observed in CHARGE syndrome. For example, the effect of mutations in a trunk NCC specific enhancer ought only be observed in derivative cell types or structures containing these derivatives and interacting cells e.g. not coloboma, atresia of the chonae, or ear abnormalities but perhaps genitourinary aberrations and potentially growth retardation might be expected.

Chapter 5

Placing a Novel NC-specific Chd7 Enhancer Element within the NC-GRN

5.1 Introduction

In this chapter I further characterised the novel Chd7 cranial NC-specific enhancer Chd7_a, identified in Chapter 4, by employing the advanced toolkit we previously developed, presented in Chapter 3, to place Chd7_a within the NC-GRN.

Chd7_a was identified as a putative enhancer by comparing dynamic NC-specific ATAC-seq profiles at HH4 and HH10-11/9-12 somite stage (R. Williams & Sauka-Spengler, unpublished data). Chd7_a appeared as a distinct peak of open chromatin in the later stage ATAC-seq profile, approximately 17kb upstream of the Chd7 gene body (see Figure 4.2.). As the region was within the topologically associated domain (TAD) of Chd7 (R. Williams & Sauka-Spengler unpublished) it was assumed to regulate Chd7. The 2.1kb putative enhancer region was cloned upstream of a ubiquitous minimal promotor sequence and fluorescent reporter and tested for enhancing activity by electroporation into HH4 chicken embryos. Fluorescent reporter signal was observed in cranial NCCs at the 9-12 somite stage (see Figure 4.4.). Given the reporter construct contained a ubiquitous minimal promoter sequence, and the

entire embryo was electroporated with the construct, the NC specific fluorescent reporter activity suggested that tissue specific factors were utilising this enhancer.

Resolving a GRN necessitates establishing the specific interactions between TF inputs and downstream output genes in order to bring order and hierarchy to the mass of interactions being summarised. This is achieved by establishing the TFs binding at the enhancer in addition to assessing the effect of these binding events on downstream gene expression.

By combining TF binding site prediction software and ChIP-based assays (-seq or -QPCR) it is possible to assay and confirm TF binding at a given enhancer. Targeted mutagenesis of the predicted binding site within the context of an enhancer and knockdown of the putative upstream TFs can be used to characterise the regulation of the downstream gene target, mediated by the given enhancer under study.

Thus, I established the essential regulatory core of Chd7_a was an 850bp region, a smaller region more amenable for TF binding site prediction and subsequent interrogation. I employed the online JASPAR database to predict TF binding motifs within Chd7_a (Sandelin et al., 2004; Vlieghe et al., 2006), and employed our previously developed toolkit to interrogate the putative upstream regulators. I mutated a number of predicted TF binding motifs within the core enhancer to determine if TF binding was necessary for Chd7_a enhancing activity, and subsequently employed CRISPR-Cas9 technology to validate that upstream TF knockout recapitulates the effect of binding site mutations on Chd7_a enhancing activity.

Additionally, I decommissioned the endogenous Chd7_a enhancer using dCas9-KRAB to induce localised chromatin compaction. By measuring Chd7 expression levels after perturbation of the chromatin landscape in the proximity of Chd7_a, this epigenome engineering approach tackled the functional role of this enhancer and its significance to the Chd7 gene locus. On account of wanting to fit Chd7_a into the NC- GRN, I focussed on FoxD3 and Sox10 biding motifs as they are known classical NC TFs.

5.2 Results

5.2.1 Chd7_a enhancer activity overlaps with that of Sox10 enhancer Sox10E2

Similar to comparing the *in situ* patterns for Chd7 to those of NC specific key TFs, FoxD3 and Sox10 (see Figure 4.1.), I also tested the co-activity of NC-specific Sox10 enhancer Sox10E2 (Betancur et al., 2011) and Chd7_a (see Figure 5.1.). I observed a very high level of activity overlap between these two enhancers in the NCCs (n=9/9) (white arrows in Figure 5-1), though most of the activity was limited to the cranial NC (panel A' in Figure 5.1.). I also noted co-expression in the otic placodes (white arrows in panel A in Figure 5.1.). Sox10E2-driven reporter expression in the anterior neural folds was not observed in the case of Chd7_a (yellow arrows in panel A' in Figure 5.1.).

5.2.2 A highly accessible 850bp region contains the essential regulatory core of the Chd7_a enhancer element

Establishing the regulatory core of the Chd7_a enhancer was necessary to use the JASPAR online utility effectively, and predict putative upstream regulators. The 2.1kb Chd7_a enhancer region lies between the histone modification peaks (H3K27AC, H3K4me1; R. Williams, data not shown). However, the ATAC-seq profile shows a single distinct region of open chromatin within this 2.1kb region (Chapter 4 Figure 2). ATAC-seq peaks indicate areas amenable to binding by TFs and can even serve to footprint TF binding sites in certain conditions (Buenrostro et al., 2013). I designed primers to amplify the 852bp fragment delimited by the ATAC-seq peak. This fragment was cloned into the fluorescent reporter vector, and its activity was assayed in the embryo in direct comparison to the initial 2.1kb element. The enhancing activity of 852bp element recapitulates that of the long enhancer entirely, both in its specific restriction to the NC and in the matching expression levels (n=35/35). Using this

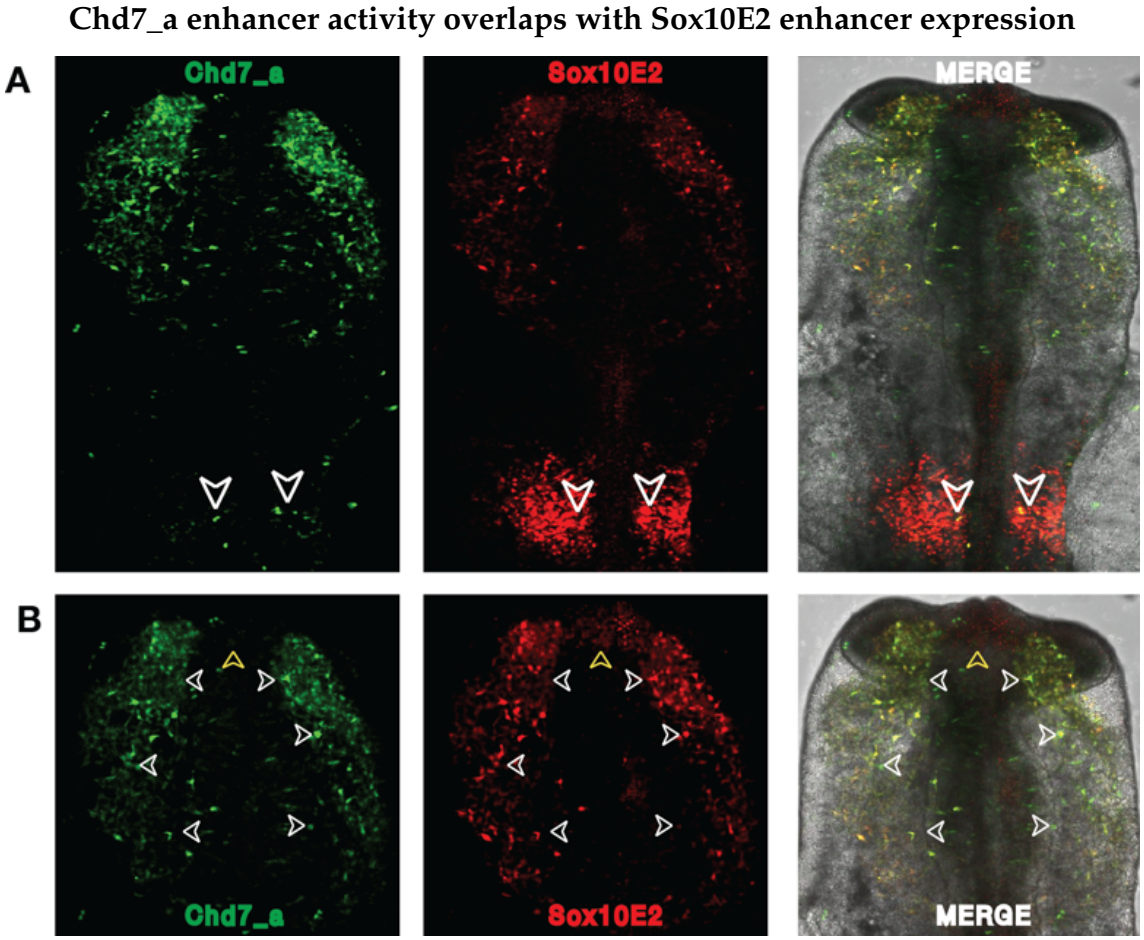


FIGURE 5.1: (A) Cranial region of a HH10 chicken embryo after co-electroporation with the Chd7_a_(green) and Sox10E2 (red) enhancer reporter expression constructs, showing regions of overlapping expression. A number of cells co-expressed the Chd7_a and Sox10E2 reporters in the otic placode. (white arrowheads). (B) Magnified image of same embryo as seen in Panel A. Chd7_a and Sox10E2 reporter expression overlapped significantly in the migrating NC cells surrounding the forebrain (white arrowheads). Chd7_a reporter expression did not extend to the anterior neural folds, while Sox10E2 reporter expression did (yellow arrowhead).

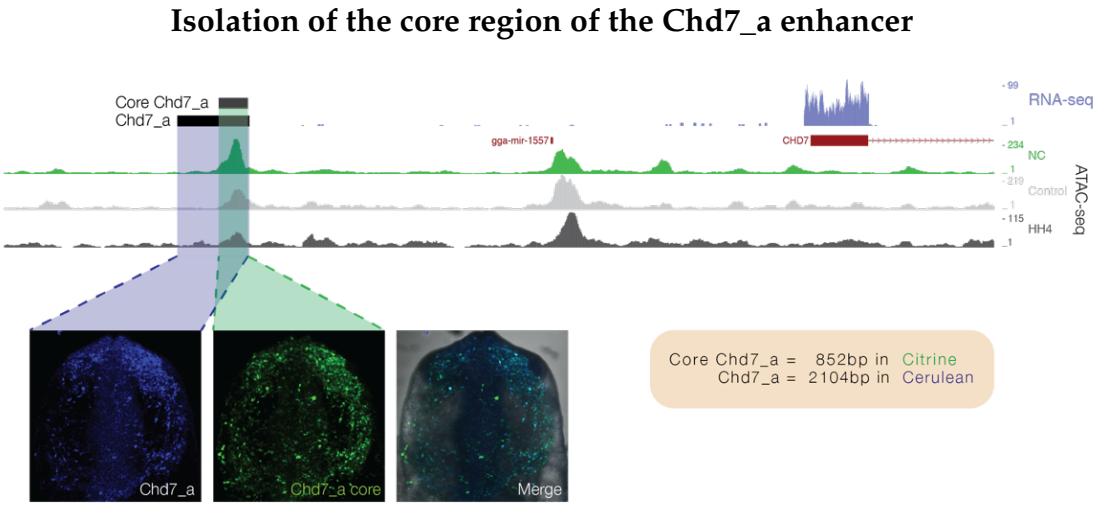


FIGURE 5.2: Schematic showing the genomic region immediately upstream of the Chd7 locus in the chicken genome. RNA-seq data (blue track), from NC1-GFP sorted NC cells around EMT, shows gene expression for the first exon of Chd7. ATAC-seq profiles show the regions of open chromatin in this region in 8-10ss NC1-GFP sorted NC cells (green), GFP-negative control cells at the same stage (grey) and cells from HH4 stage embryos (black). Black bars indicate the location of the Chd7_a putative enhancer. The initial 2.1kb Chd7_a enhancer (bottom black bar) was reduced to a 'core' 850bp region which encompassed the large ATAC-seq peak present at this location. Bottom images show a high-magnification view of the cranial region of a chicken embryo at HH10 which was co-electroporated with reporter expression constructs for both the full length (blue) and 'core' region CHd7_a enhancers (green). The reporter expression profile of the 'core' enhancer recapitulated the reporter expression from the full-length Chd7_a enhancer.

approach, I was able to reduce Chd7_a to 852bp without loss of specific activity (see Figure 5-2). Henceforth, Chd7_a refers to this refined enhancer element.

5.2.3 Avi-FoxD3 and Avi-Sox10 bind novel enhancer Chd7_a

Having refined the Chd7_a to the 852bp enhancer element, I employed the JASPAR 2014 online searchable database to predict upstream TF that can putatively bind to this region (all available vertebrate TF matrices were screened for at 87% homology, http://jaspar2014.genereg.net/). A large number of Sox and Fox family binding motifs were identified, organised in clustered hotspots (Figure 5.3., panel A) (see Appendix 6 for annotated Chd7_a sequence). I had previously demonstrated Chd7 expression correlates with that of FoxD3 and Sox10 *in vivo* (see Figure 4.1.)

and shown Chd7_a co-expression with known NC-specific FoxD3 and Sox10 enhancers (see Figure 4-4, Figure 5.1.). The FoxD3 motif used by JASPAR 2014 is 5'-A[AT]T[AG]TTTGTTT-3', whilst the Sox10 motif is 5'- (A/T)(A/T) CAA(A/T)G -3'. To test which family of TFs was of greater significance to the activity of Chd7_-a, I generated a 'No Fox/Sox Only' (Figure 5.3. Panel B) and 'No Sox/Fox only' (Figure 5.3. Panel C) constructs lacking all Fox or Sox family TF predicted binding sites, while retained all Sox or all Fox family TF sites, respectively. Each site was mutated using a single-nucleotide replacement, thus preserving the size of the enhancer, and only affecting the predicted site. I confirmed that all mutation constructs lacked the targeted sites, and that introduced mutations did not result in generation of new TF binding as predicted by the JASPAR tool again (all available TF matrices were screened for at 87% homology). Bilateral electroporation comparing the 'No Fox/Sox Only' mutated construct, where all Fox family TF predicted binding sites were mutated, to the wild- type enhancer had very mild impact on fluorescent reporter activity. The use of bilateral electroporation permits experimental and control conditions to be tested within the same embryo. However, 78% of the assayed embryos (14/18 embryos) showed specific loss of the enhancer activity in the anterior midbrain neural crest (Figure 5.3. Panel B). Conversely, I found that the 'No Sox/ Fox Only' mutations in the reporter construct, where all Sox family TF predicted binding sites were affected, had a significant impact on the Chd7_-a enhancer. Where 85% of the embryos (28/33 embryos) showed near complete loss of the fluorescent reporter, thus indicating Sox10 binding motifs are essential for Chd7_a enhancer activity (Figure 5.3. Panel C) in this context. Furthermore, I observed that there were a number of overlapping Fox and Sox family member hotspots in the central portion of the Chd7_a enhancer. I hypothesised that the two putative upstream regulators might work together in this putative regulatory core portion of the enhancer. I next generated 'Core No Sox/Fox Only' construct, which contained only single-nucleotide mutations (replacements) within the Sox binding motifs: differing from the Chd7_a 852 bp wild type region by mere 9 base pairs.

These mutations affected only Sox10 and not partially overlapping Fox motifs. This construct fully recapitulated the effect of the 'No Sox/Fox Only' construct, showing no or very minimal reporter activity (85% or 10/12 assayed embryos). This established these four Sox10 binding sites that partially overlap Fox binding motifs, situated within the putative essential core of the Chd7_a enhancers, as essential for Chd7_a activity (Figure 5.3. Panel D). This result highlights that the combined action of NC- specific Sox10 and FoxD3 transcriptional regulators upon the Chd7_a enhancer may be essential for enhancing and driving the expression of Chd7 gene, specifically in the migrating cranial NC and olfactory placode/vesicle.

5.2.4 Sox10 binding is essential for Chd7_a enhancing activity

Using 'biotagging' approach in the chicken embryo, I tested direct binding of Sox10 and FoxD3 TFs to the endogenous Chd7_a enhancer in vivo using biotin-ChIP-QPCR. Embryos were co-electroporated with specific Avi-tagged FoxD3 and Sox10 constructs expressed endogenous-like fashion, under the control of cognate enhancers (R.Williams, unpublished) and ubiquitously driven BirA. Cranial regions of ten embryos were harvested at 9-12 somites stage and pooled prior to crosslinking. After immunoprecipitation using streptavidin beads, biotin-ChIP-QPCR was used to probe for enrichment at various putative enhancer regions. QPCR was employed rather than sequencing as gathering sufficient material for a sequencing run would have been laborious and global binding profiles were not necessary. The values presented in the graph are an average of three FoxD3 and four Sox10 independent biotin-ChIP replicates.

There were many putative Sox10 sites within the predicted SOX hotspots throughout the Chd7_a enhancer, however there was only one putative FoxD3 site in the putative central core region of Chd7_a. The QPCR primers used flanked this predicted FoxD3 site (Figure 5.4. Panel A) (see Appendix 2 for further details). Our results show significant enrichment of both Avi-FoxD3 and Avi-Sox10 proteins (expressed

Analysis of putative TF binding sites within the Chd7_a enhancer

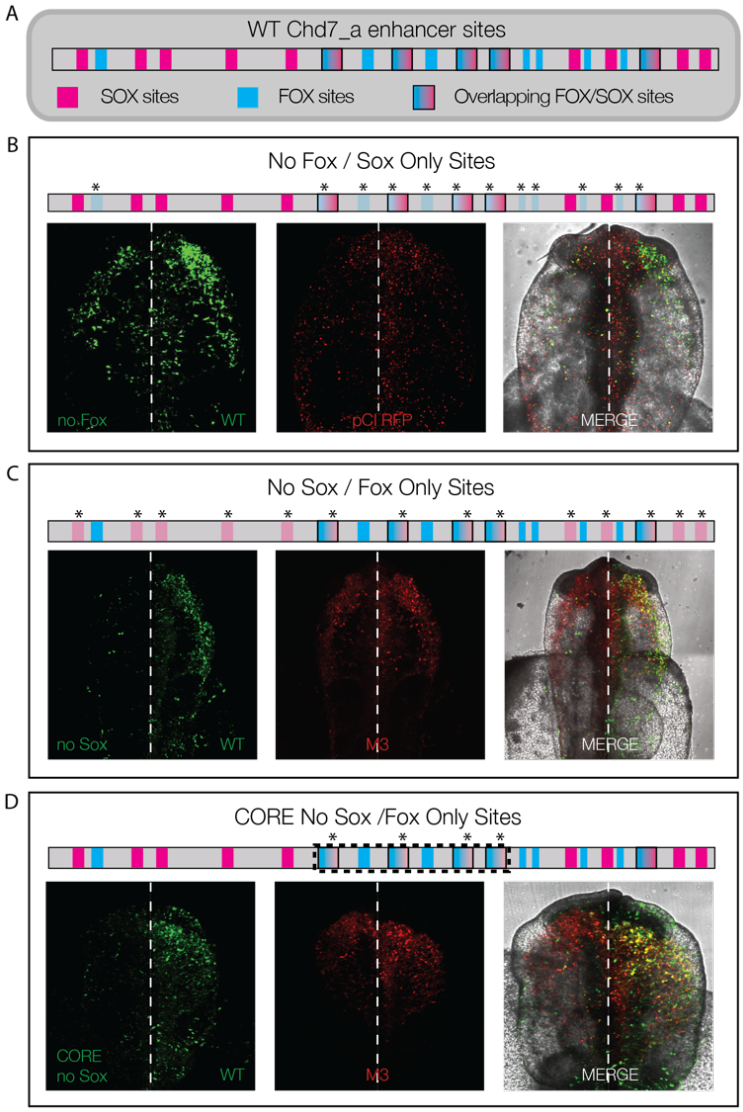


FIGURE 5.3: (A-D) Schematics of the Chd7_a enhancer indicating putative FOX (blue) and SOX (magenta) family TF binding hotspots identified using the JASPAR online utility (matrixes screened at 87% homology). Asterisks indicate the sites that were mutated by single-nucleotide replacement in each experiment, resulting in predicted binding site ablation. Images show a high-magnification view of the cranial region of chicken embryos at HH10 which were co-electroporated with reporter expression constructs for the mutated (green, left-side) and WT (green, right-side) enhancers, as well as an M3 reporter expression construct to mark NC cells (red). (B) Ablation of all putative FOX family binding sites resulted in a minor reduction in fluorescent reporter activity when the mutated enhancer activity was compared to the wild- type enhancer. (C) Ablation of all putative SOX family binding sites resulted in a major reduction in fluorescent reporter activity when the mutated enhancer activity was compared to the wild-type enhancer. (D) The essential SOX sites were located in a core region that contains many overlapping FOX and SOX binding sites (dotted-box). Ablation of these binding sites alone recapitulated the reporter expression pattern seen when all SOX sites were deleted.

as %Input) at Chd7_a genomic region, as compared to other putative NC enhancer regions tested CR3, Lmo4, Pax3 etc (5 fold higher) and compared to the negative control region (>20 fold higher) (Figure 5.4. Panel B, C). In conclusion, I find that the endogenous Chd7_a regulatory element is occupied by biotinylated Avi-tagged FoxD3 and Sox10 proteins in the cranial NC of HH10-11 chicken embryos.

Next, I tested whether Sox10 TF is required for exogenous Chd7_a activity in the migrating cranial NCCs. To this end I used transient genome engineering (CRISPR/Cas9 system) to knockout Sox10 gene and assayed the effect of this perturbation on the activity of Chd7_a fluorescent reporter. I designed two sgRNAs targeting the HMG box in the second coding exon of the Sox10 gene (Figure 5.5. Panel A) and confirmed sgRNA efficiency using HRMA and T7 assays (see Chapter 2 for further details). I employed the bilateral electroporation technique to co- electroporated the pool of two sgRNAs in U6_3 mini-vectors with ubiquitous construct expressing Cas9 into the experimental side. On the right, control side, I electroporated scrambled sgRNA + Cas9- Citrine construct (Figure 5.5. Panel B). Each side of the embryo also received Chd7_a Cerulean reporter plasmid, to allow direct assessment of the Sox10 knockout on the reporter activity. Assuming the sgRNA appropriately targeted the Cas9 activity to the Sox10 genomic locus, the subsequent Cas9 activity should result in mutation of the gene. The reporter activity on the experimental side, lacking Sox10 due to highly efficient genome engineering was nearly completely lost in 6/6 embryos assayed. When the reporter construct was co-electroporated with scrambled sgRNA on both sides, Cerulean reporter activity was unaffected. I have shown that loss of Sox sites in the Chd7_a enhancer results in a major loss of Chd7_a enhancing activity and that the 4 Sox binding sites in the core region are of particular significance. I have also demonstrated that Sox10 directly binds endogenous Chd7_-a enhancer *in vivo*. Finally, using CRISPR/Cas9-mediated Sox10 knockout in vivo, I demonstrated that Sox10 is required for proper expression of fluorescent reporter under the control of Chd7_a element. Taken together, these results show that key NC-specific TF Sox10 is essential for the normal Chd7_a regulatory activity, and

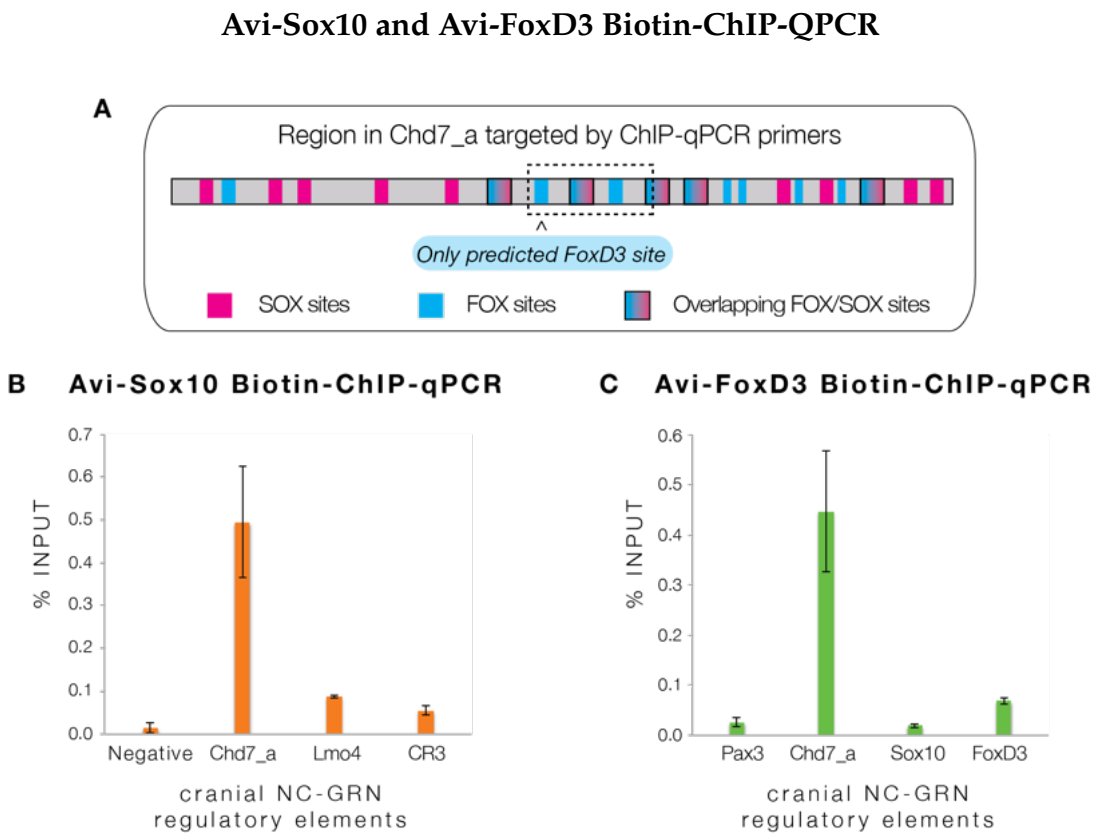


FIGURE 5.4: (A-B) Each biotin-ChIP was conducted on material derived from 6 dissected cranial regions, therefore containing approximately 5,000 NC cells of interest. All results are reported as % Input. (A) Schematic of Chd7_a showing the region targeted for ChIP-QPCR amplification in relation to the only JASPAR predicted FoxD3 binding site. (B) Avi-Sox10 pull-down resulted in an enrichment of the Chd7_a enhancer region, compared to a negative control region located on chromosome 4 in a gene desert region. Ets1 and Lmo4 are important NC specifier genes; 'CR3' primers targeted a known Ets1 enhancer and 'Lmo4' targeted a novel Lmo4 enhancer (R.Williams, unpublished) (n = 4 independent ChiP experiments). (C) Avi-FoxD3 pull-down resulted in an enrichment of the Chd7_a enhancer region. 'Pax3' primers targeted a novel Pax3 enhancer (R.Williams, unpublished), 'Sox10' primers targeted the known Sox10E2 enhancer and 'FoxD3' primers targeted the known FoxD3- NC1 enhancer. (n = 3 independent ChIP experiments).

most likely enhances tissue-specific expression of Chd7 in the NCCs.

5.2.5 Targeted epigenomic modification of Chd7_a chromatin context affects Chd7 expression levels

To establish the importance of Chd7_a enhancer for endogenous Chd7 gene expression I employed epigenome engineering approach targeting dCas9-KRAB fusion protein to the Chd7_a enhancer. KRAB activity has been shown to facilitate H3K4me1 and H3K27Ac removal, and H3K9me3 deposition, at KRAB targeted genomic regions (Groner et al., 2010; Thakore et al., 2015). The KRAB induced chromatin compaction at the Chd7_a enhancer was used to assay the significance of this enhancer on Chd7 expression. I hypothesised that interfering with chromatin context of the Chd7_a enhancer will affect its regulatory activity, resulting in a reduction in Chd7 expression levels, thus demonstrating that Chd7_a enhancer is required for proper Chd7 expression (Figure 5-6 Panel A). In the experimental conditions, I used five guide RNAs to target the chicken beta actin promoter-driven dCas9-KRAB to the Chd7_a enhancer (see Appendix 3). I confirmed successful bilateral co-electroporations using a ubiquitous reporter construct, pCI- H2B-RFP/EGFP respectively (Figure 5-6 Panel B). The use of bilateral electroporation permits experimental and control conditions to be tested within the same embryo. Chd7 is expressed throughout the cranial region at this stage. Therefore, to increase the likelihood of detecting changes in the expression of Chd7 in the NC, I dissected the dorsal aspect of the respective sides of the cranial region in analysed embryos, hence enriching for cranial NCCs. I observed a consistent down-regulation of Chd7 on the experimental side of individual embryos where specific sgRNAs were targeting dCas9-KRAB to Chd7_a compared to control scramble sgRNA conditions on the control side (Figure 5-6 Panel C, n=9/12 assayed embryos, average fold change difference calculated as ratio of expression on the experimental versus control side control within individual embryos was 0.68 +/-0.05 (0.63 - 0.73). I also confirmed there were no left-right

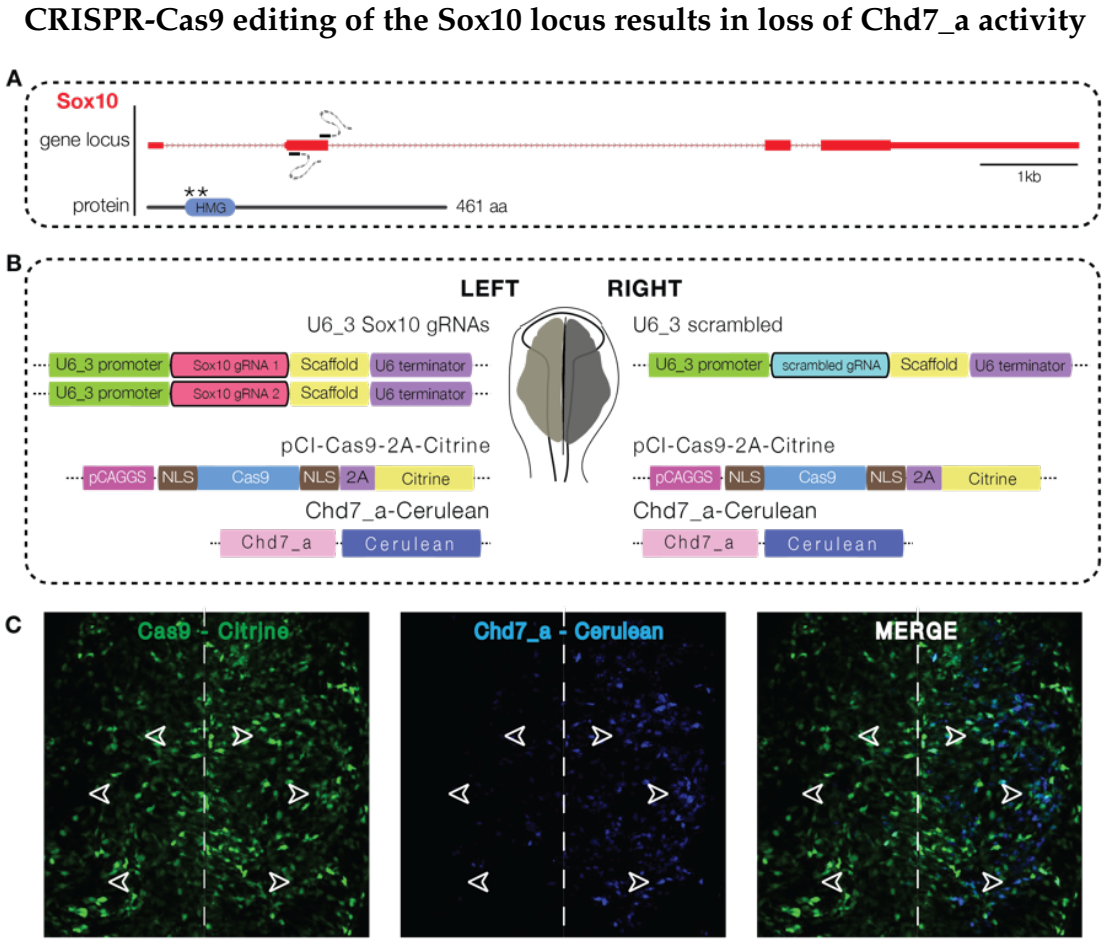


FIGURE 5.5: (A) Two gRNAs were used to target the functional HMG-box located in the second exon of Sox10. (B) Schematic showing the bilateral electroporation experimental design used. Bilateral electroporation was used to compare the effect of the two Sox10 targeting gRNAs on Chd7_a enhancer reporter activity (left side) to a scrambled control gRNA (right side). Both constructs were co-electroporated with a ubiquitous Cas9-2A-Citrine, therefore all cells containing CRISPR-Cas9 constructs are highlighted by the Citrine reporter (green). (C) High magnification view of a stage HH9 chicken embryo showing the migrating cranial NC (arrowheads). Chd7_a reporter activity (blue) in the migrating NC was lost upon CRISPR-Cas9 editing of the endogenous Sox10 locus (arrowheads indicating where Chd7_a expression is seen on the right, and should be seen on the left).

Chd7 expression bias by conducting a scrambled gRNA on both sides of the midline experiment and found no significant differences in Chd7 expression levels. Using this set of scrambled/scrambled sgRNA/dCas9-KRAB electroporated embryos as a control dataset, I found that observed effect was statistically significant, with $p<0.05$ using Chi-square test.

5.2.6 Chd7_a is syntenically conserved in human

I used DNAseI hypersensitivity, H3K4me1 and H3K27Ac ENCODE data from hESCs to identify a putative Chd7_a equivalent in the human genome (Figure 5.7. Panel A). Signal in the ENCODE data in ES cell lines, or those of NC origin, together with the position upstream of the CHD7 promotor were taken as indicative of this being a putative enhancer. There was no sequence similarity, and consequently it was not deemed as being evolutionarily conserved by ECR software. This putative enhancer, CHD7_a, lies 26kb upstream of the CHD7 start site, and was found to have enhancing activity when tested using the chicken fluorescent reporter system (Figure 5.7 Panel B, n=4/10). I observed human CHD7_a activity in chicken Chd7_a positive NCCs (Figure 5.7 Panel B' white arrows), but also noted that the activity of putative human Chd7 enhancer was broader and also found in neural tissue and ectoderm (Figure 5.7. Panel B, yellow arrows).

ENCODE data for DNAseI hypersensitivity, and enhancer-specific H3K4me1 and H3K27Ac histone marks from hESCs show that the human CHD7 locus has strong indications of a similar number of potential *cis*-regulatory elements (see Figure 5.7.). These putative human elements would also have to be tested for activity prior to establishing the potential for any clinical relevance but could constitute potential mutation harbouring sites.

Targeting of Chd7_a using dCas9-KRAB results in down-regulation of Chd7 gene expression

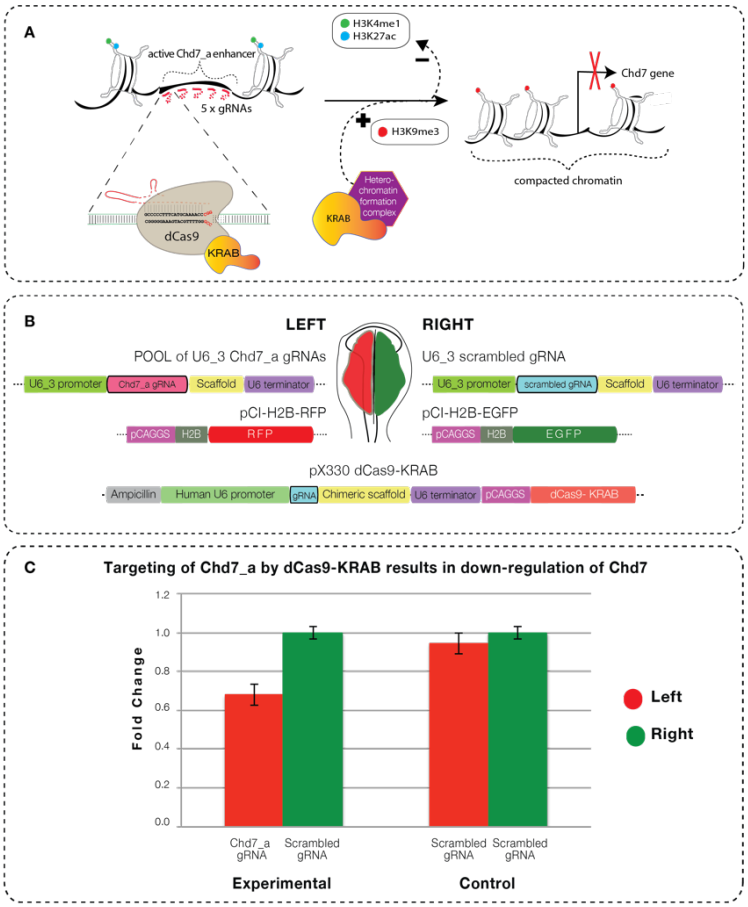


FIGURE 5.6: (A) Schematic showing the concept of dCas9-KRAB mediated chromatin compaction at the Chd7_a locus, using 5 gRNAs to direct the dCas9-KRAB to the enhancer region. dCas9-KRAB replaces the histone marks H3K4me1 and H3K27AC, which promote an open chromatin conformation, with H3K9me3, which promotes chromatin compaction. (B) Schematic showing the bilateral electroporation experimental design used for dCas9-KRAB experiments. The left side of the embryo was electroporated with a pool of 5 gRNAs targeting the Chd7_a locus, a ubiquitous H2B-RFP electroporation control reporter construct and the pX330-dCas9-KRAB construct. The right of the embryo was electroporated with a single scrambled gRNA, a ubiquitous H2B-EGFP electroporation control reporter construct and the pX330-dCas9-KRAB construct. (C) RT-QPCR on cDNA generated from HH stage8/9 embryos showed that dCas9-KRAB targeting of the Chd7_a enhancer locus resulted in down-regulation of Chd7 gene expression 0.68+/-0/05 when compared to scrambled gRNA controls (n = 9/12). Scrambled gRNAs were used to ensure the dCas9-KRAB was loaded with a gRNA but could not target any sites within the chicken genome. These results indicate there is no significant difference between Chd7 expression on either side of the midline. QPCR results are presented as fold change compared to GAPDH expression levels (ddCT method).

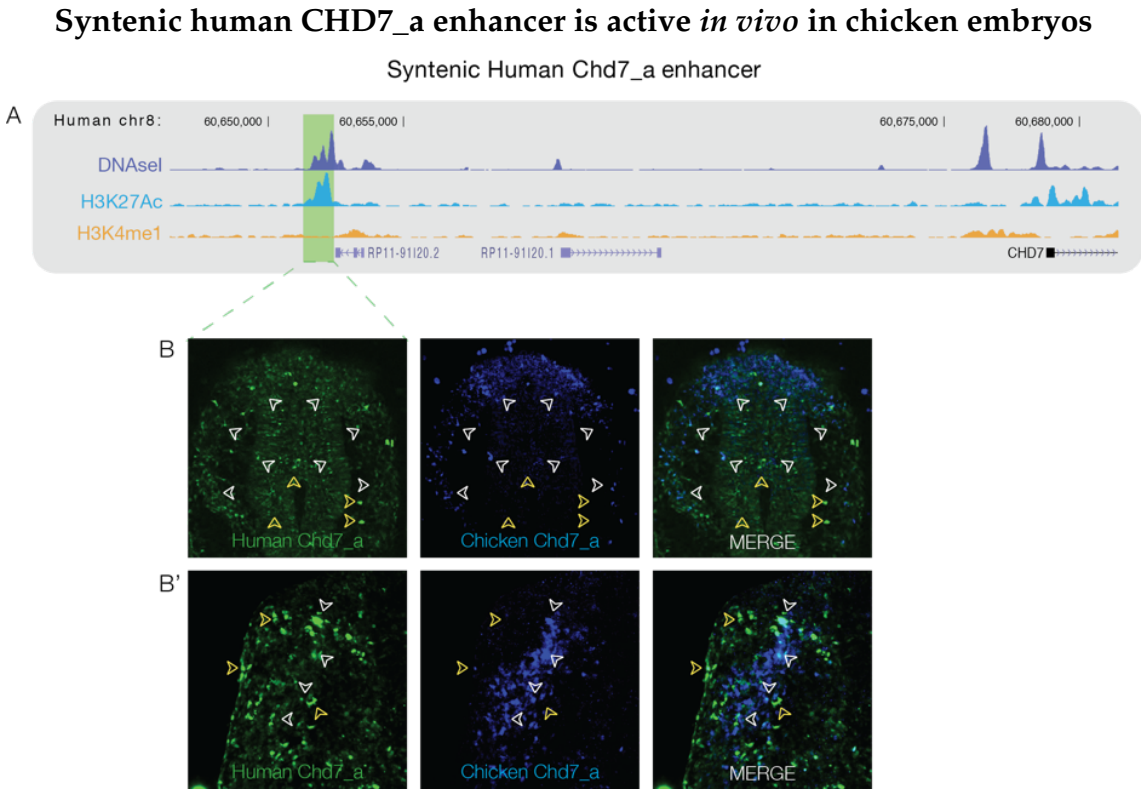


FIGURE 5.7: (A) The Human CHD7 locus with ENCODE DNAseI hypersensitivity (dark blue), H3K4me1 (orange) and H3K27Ac (light blue) tracks (Rosenbloom et al., 2013). A human CHD7_a enhancer (green highlighting), was identified using the ENCODE data and used in reporter expression assays in the chicken (B) High magnification view of a HH10 stage chicken embryo showing the migrating cranial NC. The putative human CHD7_a enhancer (green) was found to be active *in vivo*, and that the reporter expression overlapped with that of chicken Chd7_a reporter expression (blue) (white arrowheads). Human CHD7_a driven reporter expression was observed in Chd7_a negative cells also (yellow arrowheads). (B') A zoomed section and different optical slice of image shown in Panel B, showing that whilst NC-specific Chd7_a positive cells also expressed human CHD7_a reporter activity, human CHD7_a reporter activity extended beyond that of the NC (yellow arrows).

5.3 Discussion

To characterise the novel Chd7_a tissue-specific enhancer, I employed the advanced toolkit developed to probe the upstream inputs and functional activity of this enhancer. Initial work showed Chd7_a enhancer activity in M3-FoxD3 and Sox10E2 positive NC cells, establishing possible regulation by these key NC TFs. I refined the Chd7_a enhancer to a regulatory region spanning a 850bp ATAC peak within the initially cloned element (see Chapter 4 for more details). Using the JASPAR online searchable database, I predicted a number of FOX and SOX family member binding sites throughout this genomic region. Using bilateral electroporations, I tested a number of mutation constructs harbouring binding site mutations and established that Sox sites specifically in a central portion of the enhancer were essential to its activity. The use of bilateral electroporation permits experimental and control conditions to be tested within the same embryo. I then used our TF 'biotagging' technique to establish that Sox10 and FoxD3 bind the Chd7_a enhancer *in vivo*.

While Avi-tagged FoxD3 was found to bind Chd7_a by ChIP-QPCR, the Fox site mutated Chd7_a reporter construct had only marginally reduced reporter activity compared to the un-mutated wild-type enhancer reporter construct. Taken together, this suggests that while FoxD3 may bind this enhancer, FoxD3 may not be essential for efficient activity of this regulatory element in this context. It is worth noting that FoxD3, like the other members of the FOX family, is thought to act as a pioneer factor (Chen and Dent, 2014; Yong et al., 2016; Zaret and Carroll, 2011). Thus, its main effect would be increasing enhancer accessibility to other factors and complexes via modification of the local chromatin landscape. As such, the use of exogenous, non-chromatinised, constructs to test the necessity of FoxD3 binding should be reconsidered. Conversely, based on the biotin-ChIP experiments and mutation of Sox binding sites within the Chd7_a enhancer and the fact that Chd7_a enhancer activity onsets only in emigrating and migrating cranial NC, which express high levels of Sox10 gene, I concluded that Sox10 was a prime candidate, tissue-specific

TF controlling Chd7_a activity in the NC. The presence of overlapping FoxD3 and Sox10 sites could also indicate co-binding, or the necessity for FoxD3 pioneer activity prior to ensure the Sox10 sites are accessible at the appropriate time. The ChIP-QPCR confirmation of FoxD3 and Sox10 binding further strengthen the reliability of the JASPAR predictions.

I tested this hypothesis *in vivo*, by using CRISPR-Cas9 to knockout endogenous Sox10. Use of specific sgRNAs targeted to the second coding exon of the Sox10 gene and interrupting HMG binding domain, resulted in a distinct reduction of Chd7_a reporter activity. However, to further confirm that Sox10 is the necessary component in this regulatory mechanism, immunocytochemistry to show a reduction of Sox10 protein would strengthen my conclusions.

The Chd7_a enhancer lies within the Chd7 TAD, and as such it is very likely that it regulates Chd7. Epigenomic targeting of Chd7_a additionally supports this assumption, as Chd7 expression levels are reduced upon targeting of Chd7_a by dCas9-KRAB. However, the impact of this epigenome engineering was challenging to accurately assay, as Chd7 is expressed in neighbouring tissues, as well in the neural crest cells. As such, a future development could include FACS to isolate the NC cells, and then compare them to non-NC control cells. This approach would rely on a fluorescent reporter that lies upstream of the enhancer in question, as it must be expressed at equal quantities under both experimental, and control conditions. It is also possible that other identified Chd7 enhancers have a complementary activity in controlling Chd7 gene. Therefore, epigenome targeting of several NC-specific Chd7 enhancers may result in strong effect (decrease) of Chd7 gene expression. Moreover, in relation to the other Chd7 enhancers identified, it would be interesting to verify whether they are also bound by Sox10 or FoxD3 TFs, and test whether these factors are essential for their activation using Crispr/Cas9 mediated knockout.

Finally, using synteny analysis, I identified a human equivalent of the Chd7_a enhancer, which was active *in vivo*. While the novel human enhancer does not have

the same DNA sequence as that of the chicken Chd7_a, it had a DNASeI and methylation status expected of an enhancer, and like the chicken Chd7_a, lies upstream of the CHD7. I found this syntenic enhancer had activity in Chd7_a positive cells, in addition to some other neighbouring cells, and suggest that refinement of this enhancer could isolate the apparent NC population expressing it.

In conclusion, I have established that Sox10 binding of the Chd7_a enhancer is both necessary and sufficient to drive fluorescent reporter activity. FoxD3 binding was also observed, but the importance of this direct binding is difficult to assay due to the necessity of a chromatinised state to allow its pioneering activity. It is of note that other FOX family member mutations are known to result in NC dysregulation associated clinical phenotypes, suggesting that the binding of this family of factors to Chd7_1 could still have significance (Lehmann et al., 2003). I have also successfully employed our toolkit and demonstrated its potential applications in the characterisation of enhancer elements to improve our ability to resolved GRNs using the chicken model system.

Chapter 6

Conclusion and Future Directions

CHARGE syndrome is characterised by congenital malformations of craniofacial structures that either contain significant NC contributions, or are patterned by NCCs. As such, CHARGE syndrome is a neurocristopathy (Siebert et al., 1985). Estimates of CHARGE syndrome incidence range from 0.1-1.2/10,000 births worldwide (Blake and Prasad, 2006; Jongmans et al., 2006) to 1-2.8/10,000 births in Australian populations (Hsu et al., 2014a). The majority of CHARGE patients have causative mutations within the ubiquitously expressed chromatin remodeller CHD7 (Nestorov et al., 2015; Vissers et al., 2004; Zentner et al., 2010). Whilst estimates vary from approximately 67-75% of patients carrying CHD7 mutations (Blake and Prasad, 2006; Zentner et al., 2010) to 58% of cases (Lalani et al., 2006), the remaining 25-40% of cases could be attributable to mutations in downstream targets of CHD7 activity. Thus, an approach to better understand CHARGE syndrome is to look for downstream effectors of CHD7. For example, CHD7 KO models have implicated NCC guidance factors such as SEMA3A in CHARGE syndrome pathogenesis (Bartsch et al., 2018). SEMA3A is also known to be mutated in Kallmann syndrome patients, who are phenotypically highly similar to those with CHARGE syndrome (Raivio, 2014). However, CHD7 dysregulation disproportionately affects the NC, despite its ubiquitous expression during early development, indicating the characteristic presence of enhancer-based spatio-temporal control (Shlyueva et al., 2014). Therefore, the remaining 25-40% of CHARGE syndrome patients could also be attributable to

mutations in upstream *cis*-regulatory regions to CHD7 (Bergman et al., 2011; Zentner et al., 2010). For example, certain cases of Pierre Robin Sequence (PRS), characterised by NC-derived cleft palate disruptions, are attributable to the disruption of very long range *cis*-regulatory elements near SOX9 (Gordon et al., 2011; Benko et al., 2009). Additionally, identifying the TFs binding such putative NC-specific CHD7 enhancers would allow integration of CHD7 in the NC-EMT-GRN and provide insight into why only certain NC derivatives are affected in CHARGE syndrome. Deciphering GRN dynamics during normal embryonic development serves to further our understanding of the implications of causative mutations in disease, paving the way for effective implementation of these new technologies for therapeutic purposes. This work has implications for the identification of novel potential causative mutations for CHARGE-syndrome humans, in addition to showing that ubiquitous chromatin remodelling machinery can be tissue-specifically regulated via enhancer elements and key lineage factors.

Consequently, we have identified a number of putative NC-specific enhancers for this otherwise ubiquitous chromatin remodeller and developed the necessary tools to characterise this enhancer-based regulation in a tissue specific manner. Using CHD7 as a case study, I investigate the role that tissue specific regulation of a ubiquitously expressed chromatin remodeller plays in NC development and disease.

This hypothesis has contentious ramifications. The classic model that TFs drive differentiation and that chromatin regulators are primarily responsible for simply stabilising differentiated states is being questioned. Increasing evidence implicates chromatin regulators in priming transcriptional responses prior to cell fate decisions, modulating gene expression during differentiation and establishment of cell identity, and maintaining epigenetic marks in differentiated cells (Chen and Dent, 2014; Nestorov et al., 2015). For example, enhancers are pre-patterned by H3K4me1 / H3K4me2 marks and histone variants H3.3/H2A.Z prior to activation of the associated gene. The presence of H3K4me1 and H3K4me2 marks may facilitate the

binding of pioneer transcription factors (Zaret and Carroll, 2011). Alternatively, pioneer TFs access closed chromatin, facilitate the recruitment of alternative histones to these sites, and generate permissive chromatin states for the action of tissue-specific TFs (Golson and Kaestner, 2016). Interestingly, the Fox family of TFs are known to have pioneer factor activity on account of their structural similarity to linker histones H1 and H5 (Zaret and Carroll, 2011). Additionally, enhancers are employed frequently in development to encode spatio-temporal information.

While the importance of epigenetic pre-patterning in spatio-temporal regulation of developmental gene expression via enhancers is undoubtedly important, it is also a widespread phenomenon in cell fate decisions (Chen and Dent, 2014) and consequently observed during somatic cell reprogramming into induced pluripotent stem cells (iPSCs). Circular chromosome conformation capture with high throughput sequencing (4C-seq) found establishment of long-range interactions to the Oct4 and Nanog loci precedes the transcriptional activation necessary for successful iPSC reprogramming(Apostolou et al., 2013). This is noteworthy as it indicates that these long- range interactions may be essential to the subsequent gene activation and could be used to identify potential upstream regulators.

To investigate this contentious possibility of NC-specific enhancers for CHD7, as a ubiquitously expressed chromatin remodeller, I employed the classic NC and enhancer- screening model system; the chicken (reviewed by Stern (2005)). Cranial NCCs were specifically FACS isolated using a cranial NC FoxD3 enhancer reporter, NC1-GFP. These cranial NCCs were then profiled using ATAC-seq to highlight areas of high accessibility, commonly associated with regulatory elements, including enhancers. It was possible to employ classical techniques to screen putative, ATAC-seq highlighted, Chd7 enhancers using fluorescent reporter constructs electroporated into chicken embryos. Ten putative cranial NC specific enhancers were identified using this approach. However, the deeper characterisation of novel NC-specific enhancers required the translation of TF biotagging and CRIPSR-Cas9 genome- and

epigenome- engineering techniques for use in the chicken system. Having translated these advanced techniques for use in the chicken, I employed them to interrogate one of the identified novel cranial NC specific Chd7 enhancers, Chd7_a.

The limited number of NCCs in the chicken embryo required the development of an advanced toolkit of techniques to explore gene regulatory interactions under such conditions. I adapted the 'biotagging' technique, developed by Sauka-Spengler lab in zebrafish for use in profiling of transcription factor binding in the chicken embryo (Trinh et al., 2017). Additionally, we have developed a new toolkit, optimising the CRISPR-Cas9 technologies in chicken embryos, such that it can provide targeted genomic and epigenomic modifications in somatic embryonic cells with high efficiency (Williams et al., 2018). By comparing the chromatin accessibility profiles of non- NCCs to NCCs before and after their obligatory EMT. I identified seven novel Chd7 enhancers, including two strong cranial neural crest-specific enhancers. Other novel Chd7 enhancers exhibited tissue-specific activity in trunk or vagal NCCs, but also an enhancer with apparent ubiquitous activity was also identified. Utilising the novel methodologies developed for characterisation of gene regulatory interactions in the chicken embryo, I probed the novel cranial NC-specific Chd7 enhancer Chd7_a and demonstrated that its activity is required for proper expression of Chd7 in the neural crest. Furthermore, by mutating predicted TF sites and observing the loss of reporter activity, I demonstrated that NC-specific TF Sox10 is the key upstream input into the novel Chd7_a enhancer. Additionally, by using a Cas9-KRAB to alter the chromatin context surrounding Chd7_a, I showed endogenous Chd7 transcription was reduced, indicating the significance of Chd7_a enhancing activity on Chd7 expression *in vivo*.

Whilst the TF-biotagging tools developed here will facilitate the dissection of other important GRNs using the chicken system, the CRISPR-Cas9 tools are particularly timely given recent advances in the generation of transgenic chickens. The *in vitro* editing of chicken primordial germ cells (PGCs), and subsequent reintroduction into host chickens harbouring mutations enhancing the transmission rate

of edited PGCs, in particular stands to benefit from the advanced toolkit presented here (Bednarczyk et al., 2018; Nishijima and Iijima, 2012; Taylor et al., 2017). Using this advanced toolkit, I demonstrated that reduced accessibility of Chd7_a impacts Chd7 expression levels, indicating Chd7_a enhances the expression of Chd7. Using biotagged FoxD3 and Sox10 constructs, I demonstrated these key NC TFs bind Chd7_a in vivo. Using reporter constructs with mutations in various combinations of predicted Fox and Sox family binding sites, I demonstrated that Sox10 binding, in particular, is essential to Chd7_a enhancing activity. I confirmed the necessity of Sox10 for Chd7_a activity by targeting Sox10 with CRISPR-Cas9 and demonstrating a loss of Chd7_a reporter construct activity. Finally, I showed that Chd7_a has a syntenic human enhancer, which has activity in the chicken NC.

However, beyond the development of an advanced toolkit for GRN dissection using the chicken, there are a number of potential implications to the finding of NC-specific Chd7 regulation. Most directly, my work places Chd7 within the NC-GRN, under the direct regulation of Sox10. The potential for FoxD3 regulating CHD7 via Chd7_a must not be underestimated, given mutations in other FOX family members, FOXE1/3, FOXC1, and FOXL2 are associated with human syndromes featuring NC derivative aberrations including glaucoma, ocular anterior segment formation, cleft palate and drooping eyelids (Lehmann et al., 2003). More broadly, this work highlights that lineage-specific TFs can regulate chromatin machinery via enhancers. Elucidating the NC-GRN will improve our understanding of the forces regulating a cell population that is a major source of birth defects in humans.

Additionally, some of the 30% of unaccounted for cases of CHARGE syndrome patients may well harbour mutations in the human equivalent Chd7_a enhancer, indicating potential clinical significance for this novel NC specific Chd7 enhancer. Whilst simply identifying the mutations in patients suffering from CHARGE syndrome does little to aid their prognosis, recent advances in CRISPR-Cas9 and stem cell technologies hold therapeutic promise for a number of genetic diseases and

birth defects in humans (Musunuru, 2013). Deciphering GRN dynamics during normal embryonic development serves to further our understanding of the implications of causative mutations in disease, paving the way for effective implementation of these new technologies for therapeutic purposes. For example, the confirmation of human Chd7_a involvement in CHARGE could be used as a diagnostic site in IVF associated embryonic screening. If CHARGE syndrome causative mutations were identified in these putative human enhancers, embryos generated using *in vitro* fertilisation (IVF) techniques could be screened for these mutations. Looking even further into the future, it may even become possible to use CRISPR-Cas9 genome editing to rectify any pathogenic mutations found in these locations.

If this work were to be taken further, the characterisation of the other novel enhancers identified here would be important. For example, the novel enhancers found to have trunk and vagal NC specific activity should be confirmed by co-electroporation with other known trunk and vagal NC markers where possible (Chan et al., 2005; Simões-Costa et al., 2012). Other neuronal and placodal enhancers e.g. six1 or n-cad (Betancur et al., 2011; Matsumata et al., 2005; Sato et al., 2012) should also be screened for co-expression with these novel enhancers to further establish which cell types specifically regulate Chd7 during early development. Furthermore, looking for potential synteny between the other novel chicken Chd7 enhancers, not characterised within the scope of this project, and the human Chd7 locus would be pertinent. Additionally, the conservation of Sox10 and FoxD3 binding to the human Chd7_a enhancer could be tested using hNCLCs. This would involve: looking for reporter activity for the human Chd7_a enhancer in hNCLCs, SOX10/FOXD3 ChIP-qPCR at putatively syntenic enhancer, knocking out SOX10/ FOXD3 and looking for loss of reporter activity and associated reduction of endogenous Chd7 levels. Alternatively, bioinformatics assessment of predicted TF binding sites in any of the syntenic enhancers would be highly interesting. It would also be necessary to confirm the putative human Chd7_a enhancer equivalent does directly bind to CHD7 in human using methods like Capture- C, as it is just synteny and proximity that

currently suggests that CHD7 is the target of regulation by this enhancer. *in vivo* confirmation of the necessity of any potential human CHD7 enhancers would also be necessary.

Aside from the obvious question surrounding the potential clinical significance of Chd7_a, a number of broader questions remain unanswered. For example, are other chromatin remodellers, like DNMT3A, with significance to the NC also tissue-specifically regulated? Regarding the other novel enhancers presented, but not characterised here, what TFs bind them? How do these presumably differing combinations of TFs compare to the TFs binding the ubiquitous enhancer also presented within this work? Might it be possible to elucidate some higher organisational mechanism regarding chromatin remodelling in general from characterising this ubiquitous Chd7 enhancer? Another ubiquitous role of CHD7 is its involvement in the transcription of ribosomal DNA genes (Zentner et al., 2010). It might be interesting to interrogate the apparent dependency on CHD7 in certain tissues in terms of their respective necessity for ribosomes. Perhaps the absence of CHD7 is most apparent at times of transcriptional hyperactivity such as times of cell-fate decision and migration?

It is striking that whilst CHD7 has evident significance to the NC, this significance is not apparent in all NC derivatives. The cranial and cardiac NC appear particularly sensitive to loss of CHD7 activity, which begs the question- why? Finally, beyond CHD7 and CHARGE syndrome, it seems that NC differentiation is dependent on a highly delicate balance of chromatin regulators. This sensitivity is highlighted by the eponymous existence of craniofacial defects in human diseases with underlying defects in chromatin structure and modification (Hendrich and Bickmore, 2001). Why this is so remains a key open question.

Appendix A

Putative Chd7 enhancers

TABLE A.1: Putative Chd7 Enhancers

Putative Chd7 Enhancer Name	Distance to Chd7 TSS (Kb)	Chromosome Location	Size (bp)
Chd7_a	-17.3	chr2:112,594, 568-112,596,672	2104
Chd7_b	35.9	chr2:112,648,524-112,649,219	230
Chd7_c	75.1	chr2:112,687,882-112,688,261	415
Chd7_d	92	chr2:112,704,220-112,705,835	2356
Chd7_e	98.5	chr2:112,711,257-112,711,963	230
Chd7_f	115.3	chr2:112,726,796-112,729,745	230
Chd7_g	137.7	chr2:112,750,489-112,750,902	230
Chd7_h	153.3	chr2:112,765,767-112,766,759	230
Chd7_i	154.9	chr2:112,767,559-112,768,364	230
Chd7_j	166	chr2:112,778,855-112,779,372	556

TABLE A.2: Putative enhancer primers for GoldenGate assembly

Primer Name	Primer Sequence
Chd7_a_fwd	TTTTTTCGTCTCGCCAGGCCTTACTTTTCACGACAGAAGCT
Chd7_a_rev	TTTTTTCGTCTCCAACAGCTCGCCACGTTTCTCAGATG
Chd7_b_fwd	TTTTTTCGTCTCGCCAGGGGCAGTGGTAGAGTACGTGT
Chd7_b_rev	TTTTTTCGTCTCCAACAGCTGTCTGCTTGGTCCCCTAT
Chd7_c_fwd	TTTTTTCGTCTCGCCAGGCTGCAGCCCACCTTAAAACC
Chd7_c_rev	TTTTTTCGTCTCCAACAGGGCTTTGTGTCTGACCCATG
Chd7_d_fwd	TTTTTTCGTCTCGCCAGGAATCCCCGTCCCTCTTTTCC
Chd7_d_rev	TTTTTTCGTTCTCAACAGAGATGCACAAATCCCCTTGC
Chd7_e_fwd	TTTTTTCGTGCCAGGCAGGACAAGGCAGCATCTTC
Chd7_e_rev	TTTTTTCGTCTCCAACAGATTTCTGGTGCCTGCTGATG
Chd7_f_fwd	TTTTTTCGTCTCGCCAGGAAGTCTGTGAGCATCCTGGT
Chd7_f_rev	TTTTTTCGTCTCCAACAGTTGGTGGCACGAATGAAAGG
Chd7_g_fwd	TTTTTTCGTCTCGCCAGGCTGCTCCCCATACAATGCAC
Chd7_g_rev	TTTTTTCGTCTCCAACAGGTCACTTGCCATGTATGTCCA
Chd7_h_fwd	TTTTTTCGTCTCGCCAGGAGCCTGGAATCCTGCCATAC
Chd7_h_rev	TTTTTTCGTCAACAGAGATAGGGAGATGGGTGGGT
Chd7_f_rev2	TTTTTTCGTCAACAGAGATAGGGAGATGGGTGGGT
Chd7_g_rev2	TTTTTTCGTCTCCAACAGTCTGCTGGGTCTTGGAGAAG

Appendix B

QPCR primer details

TABLE B.1: ChIP-QPCR Primers

Primer Name	Primer Sequence
CR3 fw	GCCTCCCAGTCAGTCTCATTG
CR3 rev	GGATTTAGGACGGAGGCGAG
sox10E2 fw	TGATAAGCAGCAGGAGCA
Sox10E2 rev	TGAGCAGGTTGCTGTTTAAA
Negative fw	GGTTGGATTTCCAGTCTCCA
Negative rev	TGTTTTGCTGGACAATCTGC
Chd7 fw	TTGTGTCGGTGCTGTGGAAT
Chd7 rev	CTCCTCTGTGCTGATCCTGC
Lmo4 fw	TGTCCAAATTACAGCCCCAGG
Lmo4 rev	GGCCTGAGAAAAGTGTGTCCT

TABLE B.2: RT-QPCR Primers

Primer Name	Primer Sequence
Chd7_RTqPCR 1	GCTTCACCAGTCACCACAAA
Chd7_RTqPCR REV1	CCTGAGTCATATCCGGCACT

Appendix C

CRISPR toolkit component primers

TABLE C.1: CRISPR toolkit component primers

Primer Name	Primer Sequence
fw_scrambled gRNA	TTTCGTCTCCAGTCTGCAGTGCTTCAGCCGCTGTTTAGAGACGAAA
rev_scrambled gRNA	TTTCGTCTCTAAACAGCGGCTGAAGCACTGCAGACTGGAGACGAAA
Chd7_2 gRNA1F	TTCGTCAGTCGCCATGAATTAACTGCTTTGTTTAGAGACGAAA
Chd7_2gRNA1R	TTTCGTCTCTAAACAAAGCAGTTAATTCATGGCGACTGGAGACGAAA
Chd7_2 gRNA2F	TTTCGTCTCCAGTCGACGCTCGGGCTATTTGTCGTTTAGAGACGAAA
Chd7_2gRNA2R	TTTCGTCTCTAAACGACAAATAGCCCGAGCGTCGCTGGAGACGAAA
Chd7_2 gRNA3F	TTTCGTCTCCAGTCGTTTTCCTAAATCACAATGGTTTAGAGACGAAA
Chd7_2gRNA3R	TTTCGTCTCTAAACCATTGTGATTTAGGAAAACGACTGGAGACGAAA
Chd7_2 gRNA4F	TTTCGTCTCCAGTCGTCACTGTAGTTTTACATTGTTTAGAGACGAAA
Chd7_2gRNA4R	TTTCGTCTCTAAACAATGTAAAACTACAGTGACGACTGGAGACGAAA
Chd7_2 gRNA5F	TTTCGTCTCCAGTCGCTGCAAGCCAGCAGGAGTGTTTAGAGACGAAA
Chd7_2gRNA5R	TTTCGTCTCTAAACACTCCTGCTGGCTTGCAGCGACTGGAGACGAAA
pCI KRAB infusFw	TTTGGCAAAGAATTGCTCGAGCCACCATGGACTACAAAGAAC
pCI KRAB infus Rev	CTGAGGAGTGAATTGCGGCCATCAACCGGTCTTGTACAGCT
infusion KRAB 3′	GTCGAGGCTGATCAGCGAGCTCTAGGA
5′ EcoRI Cterm Cas9	GGAACAACACAAACACTACCTTG
3′ EcoRI Cterm Cas9	TCTTCTTCTTGGGGTCAGCC

Appendix D

Nanostring Codeset accession numbers

TABLE D.1: Nanostring Codeset Accession Numbers

Name	Accession Number
FZD2	NM_204222.1
ID2	NM_205002.1
DLX5	NM_204159.1
WNT8b	AY753292.1
PCNA	NM_204170.1
WNT8A	NM_205531.1
WNT6	NM_001007594.1
EEF1A1	NM_204157.2
MYCN	NM_001031091.1
Otx2	NM_204520.2
SIX1	NM_001044685.1
GATA3	NM_001008444.1
Dach1	AF427472.1
Integrinb1	NM_001039254.1
SOX3	NM_204195.1
CCND1	NM_205381.1

WNT3A	NM_204675.1
Pax6	NM_205066.1
Six3	NM_204364.1
SOX8	NM_204731.1
FGF8	NM_001012767.1
RhoA	NM_204704.1
BCL2	NM_205339.1
MSX1	NM_205488.1
TFAP2A	NM_205094.1
ADAM10	NM_204261.1
FGF2	NM_205433.1
OPRL1	XM_417424.2
SOX10	NM_204792.1
FZD7	NM_204221.1
BMP4	NM_205237.2
SOX9	NM_204281.1
PAX7	NM_205065.1
BMP2	NM_204358.1
SFRP1	NM_204553.1
SNAI1	NM_205142.1
SNAI2	XM_419196.2
ZIC1	NM_204254.1
Pax2	NM_204793.1
mRFP1_coding	TSS_00002.1
Six4	AJ133778.1
Foxl	XM_001232088.1
HPRT1	NM_204848.1
Lef1	NM_205013.2

MYB	NM_205306.1
Id3	AY040528.1
FZD1	NM_001030337.1
Cad6	NM_001001758.1
Gapdh	AF047874.1
Tac-1	XM_418674.2
Cad7	NM_204187.2
IRX1	NM_001030338.1
SDHA	XM_419054.2
RhoB	NM_204909.1
EGFP_coding	TSS_00001.1
SOX2	NM_205188.1
Ecad	NM_001039258.1
Perp	BX935976.1
PAX3	NM_204269.1
Zic2	BU319620.1
TP53	NM_205264.1
GBX2	NM_205068.1
wnt-1	AY655699.1
Cx43A1	NM_204586.1
TWIST1	NM_204739.1
Rac1	NM_205017.1
cyp26A1	NM_001001129.1
MYC	NM_001030952.1
Ncad	NM_001001615.1
MSX2	NM_204559.1
Itegrina5	NM_205439.1
Eya1	AJ008002.1

TNIP1	NM_205298.2
Eya2	NM_204915.1
ERNI	NM_001080874.1
DLX3	NM_204804.1
FOXD3	NM_204951.1
LOC416710	XM_415007.2
CDKN1B	NM_204256.2
mRFP1_coding	
EGFP_coding	

Appendix E

Chicken Chd7 *in situ* probe details

TABLE E.1: Chicken Chd7 in Situ Probe Details

Primer Name	Primer Sequence
Chd7 **SP6 5′	GATTTAGGTGACACTATAGCCATCAAACAGTGGTCAAGGA
Chd7 **SP6 3′UTR	GATTTAGGTGACACTATAGTGAGGATGGCGCTTCAAAAG
Chd7 T7 5′	TAATACGACTCACTATAGGGGGGGTCTTGGGTTCCTTCTT
Chd7 T7 3′UTR	TAATACGACTCACTATAGGGGGTGAACTCCCGAACAGTC

Appendix F

Annotation of Chd7_a enhancer core region predicted sites

5′- GGCCTGTATCTTGCTTTCTTCAACCACTTGAAATTCTTTGGGAAGGG

FoxC1 Sox10

AATCTTCTAGAGGGTTGTGTCGGTGCTGTGGAATAAACAGTTGTCG

FoxD3

TGTGAAACTGATGTGAGCCTTAGTTTTACTCTGCCTCCACTAAAATA

Sox2 FoxC1

CACTTTTACGGTAGTTTGTGCAGGATCAGCACAGAGGAGTACAAAG

FoxC1 Sox10

AGGGAGCTTTTCAGTACAAAGGCT - 3′

FoxC1 Sox10

As predicted by JASPAR 2014 software (http://jaspar2014.genereg.net)

87% homology matrix settings.

Bibliography

M. S. Adams and M. Bronner-Fraser. Review: The role of neural crest cells in the endocrine system. *Endocrine Pathology*, 20(2):92–100, Jun 2009.

S. J. Agrawal and P. Y. Lwigale. *Quail-Chick Chimeras and Eye Development*, pages 255–263. Springer New York, New York, NY, 2015. ISBN 978-1-4939-1164-6.

J. D. Ahlstrom and C. A. Erickson. The neural crest epithelial-mesenchymal transition in 4D: a 'tail' of multiple non-obligatory cellular mechanisms. *Development*, 136(11):1801–1812, May 2009.

T. Akitaya and M. Bronner-Fraser. Expression of cell adhesion molecules during initiation and cessation of neural crest cell migration. *Developmental Dynamics*, 194(1):12–20, May 1992.

D. J. Anderson. The neural crest cell lineage problem: Neuropoiesis? *Neuron*, 3(1): 1–12, 1989.

E. Apostolou, F. Ferrari, R. M. Walsh, O. Bar-Nur, M. Stadtfeld, S. Cheloufi, H. T. Stuart, J. M. Polo, T. K. Ohsumi, M. L. Borowsky, P. V. Kharchenko, P. J. Park, and K. Hochedlinger. Genome-wide chromatin interactions of the Nanog locus in pluripotency, differentiation, and reprogramming. *Cell Stem Cell*, 12(6):699–712, June 2013.

M. Aramaki, T. Kimura, T. Udaka, R. Kosaki, T. Mitsuhashi, Y. Okada, T. Takahashi, and K. Kosaki. Embryonic expression profile of chicken CHD7, the ortholog of the causative gene for CHARGE syndrome. *Birth Defects Research Part A: Clinical and Molecular Teratology*, 79(1):50–57, Jan. 2007.

A. Arshad, C. Powell, and M. P. Tighe. Hirschsprung's disease. *BMJ*, 345(oct01 2): e5521–e5521, Oct. 2012.

A. Baggiolini, S. Varum, J. M. Mateos, D. Bettosini, N. John, M. Bonalli, U. Ziegler, L. Dimou, H. Clevers, R. Furrer, and L. Sommer. Premigratory and migratory neural crest cells are multipotent *in vivo*. *Cell Stem Cell*, 16(3):314–322, Mar. 2015.

R. Bajpai, D. A. Chen, A. Rada-Iglesias, J. Zhang, and Y. Xiong. CHD7 cooperates with PBAF to control multipotent neural crest formation. *Nature*, 463(7283):958–962, 2010.

C. V. Baker and M. Bronner-Fraser. The origins of the neural crest. part i: embryonic induction. *Mechanisms of Development*, 69(1):3 – 11, 1997.

E. Balaban, M. Teillet, and N. Le Douarin. Application of the quail-chick chimera system to the study of brain development and behavior. *Science*, 241(4871):1339–1342, 1988.

L. A. Barlow. Cranial nerve development: Placodal neurons ride the crest. *Current Biology*, 12(5):R171 – R173, 2002.

O. Bartsch, A. Borchers, L. Freese, J. Möller, C. Neuhofer, S. Pauli, I. Schanze, J. Schwenty-Lara, A. Tzschach, R. Ufartes, C. M. A. van Ravenswaaij-Arts, P. Wehner, and M. T. Y. Wong. Sema3a plays a role in the pathogenesis of CHARGE syndrome. *Human Molecular Genetics*, 75:1–10, Feb. 2018.

M. L. Basch, M. Bronner-Fraser, and M. I. García-Castro. Specification of the neural crest occurs during gastrulation and requires Pax7. *Nature*, 441(7090):218–222, May 2006.

M. A. Basson and C. van Ravenswaaij-Arts. Functional Insights into Chromatin Remodelling from Studies on CHARGE Syndrome. *Trends in Genetics*, 31(10):600–611, Oct. 2015.

A. G. Baynash, K. Hosoda, A. Giaid, J. A. Richardson, N. Emoto, R. E. Hammer, and M. Yanagisawa. Interaction of endothelin-3 with endothelin-B receptor is essential for development of epidermal melanocytes and enteric neurons. *Cell*, 79(7):1277–1285, Dec. 1994.

N. L. Becker R, Stiemer B and E. M. Mild ventriculomegaly, mild cerebellar hypoplasia and dysplastic choroid plexus as early prenatal signs of charge association. *Fetal Diagnosis and Therapy*, 16:280–283, 2001.

D. Beckett, E. Kovaleva, and P. J. Schatz. A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Science*, pages 1–9, Mar. 1999.

M. Bednarczyk, I. Kozłowska, P. Łakota, A. Szczerba, K. Stadnicka, and T. Kuwana. Generation of transgenic chickens by the non-viral, cell-based method: effectiveness of some elements of this strategy. *Journal of Applied Genetics*, pages 1–9, Jan. 2018.

M. Bellard, G. Dretzen, F. Bellard, J. S. Kaye, S. Pratt-Kaye, and P. Chambon. Hormonally induced alterations of chromatin structure in the polyadenylation and transcription termination regions of the chicken ovalbumin gene . *EMBO journal*, pages 1–8, May 1986.

A. Bellmeyer, J. Krase, J. Lindgren, and C. LaBonne. The Protooncogene c-Myc Is an Essential Regulator of Neural Crest Formation in Xenopus. *Developmental Cell*, 4 (6):827–839, June 2003.

R. Ben-Hamo and S. Efroni. Gene expression and network-based analysis reveals a novel role for hsa-miR-9 and drug control over the p38 network in glioblastoma multiforme progression. *Genome Medicine*, 3(11):77, Nov. 2011.

S. Benko, J. A. Fantes, J. Amiel, D.-J. Kleinjan, S. Thomas, J. Ramsay, N. Jamshidi, A. Essafi, S. Heaney, C. T. Gordon, D. McBride, C. Golzio, M. Fisher, P. Perry,

V. Abadie, C. Ayuso, M. Holder-Espinasse, N. Kilpatrick, M. M. Lees, A. Picard, I. K. Temple, P. Thomas, M.-P. Vazquez, M. Vekemans, H. R. Crollius, N. D. Hastie, A. Munnich, H. C. Etchevers, A. Pelet, P. G. Farlie, D. R. Fitzpatrick, and S. Lyonnet. Highly conserved non-coding elements on either side of SOX9 associated with Pierre Robin sequence. *Nature Genetics*, 41(3):359–364, Feb. 2009.

J. E. H. Bergman, N. Janssen, L. H. Hoefsloot, M. C. J. Jongmans, R. M. W. Hofstra, and C. M. A. van Ravenswaaij-Arts. CHD7 mutations and CHARGE syndrome: the clinical implications of an expanding phenotype. *Journal of Medical Genetics*, 48(5):334–342, May 2011.

P. Betancur, M. Bronner-Fraser, and T. Sauka-Spengler. Genomic code for Sox10 activation reveals a key regulatory enhancer for cranial neural crest. *PNAS*, 107 (8):3570–3575, Feb. 2010a.

P. Betancur, M. Bronner-Fraser, and T. Sauka-Spengler. Assembling neural crest regulatory circuits into a gene regulatory network. *Annual Review of Cell and Developmental Biology*, 26(1):581–603, 2010b.

P. Betancur, T. Sauka-Spengler, and M. Bronner. A Sox10 enhancer element common to the otic placode and neural crest is activated by tissue-specific paralogs. *Development*, 138(17):3689–3698, Sept. 2011.

S. Bhatt, R. Diaz, and P. A. Trainor. Signals and switches in mammalian neural crest cell differentiation. *Cold Spring Harbor Perspectives in Biology*, 5(2), 2013.

F. Bilan, M. Legendre, V. Charraud, B. Manière, D. Couet, B. Gilbert-Dussardier, and A. Kitzis. Complete screening of 50 patients with CHARGE syndrome for anomalies in the CHD7 gene using a denaturing high-performance liquid chromatography-based protocol: new guidelines and a proposal for routine diagnosis. *The Journal of Molecular Diagnostics*, 14(1):46–55, Jan. 2012.

B. R. Bill, A. M. Petzold, K. J. Clark, L. A. Schimmenti, and S. C. Ekker. A primer for morpholino use in zebrafish. *Zebrafish*, 6(1):69–77, Mar. 2009.

K. D. Blake and C. Prasad. Charge syndrome. *Orphanet Journal of Rare Diseases*, 34 (1):1750–1172, 2006.

K. D. Blake, S. L. Davenport, B. D. Hall, M. A. Hefner, R. A. Pagon, M. S. Williams, A. E. Lin, and J. M. Graham. CHARGE association: an update and review for the primary pediatrician. *Clinical Pediatrics*, 37(3):159–173, Mar. 1998.

D. E. Bockman and M. L. Kirby. Dependence of thymus development on derivatives of the neural crest. *Science*, 223(4635):498–500, Feb. 1984.

R. P. Bolande. Neurocristopathy: Its Growth and Development in 20 Years. *Pediatric Pathology & Laboratory Medicine : Journal of the Society for Pediatric Pathology, affiliated with the International Paediatric Pathology Association*, 17(1):1–25, Apr. 2010.

A. Bolotin, B. Quinquis, A. Sorokin, and S. D. Ehrlich. Clustered regularly interspaced short palindrome repeats (crisprs) have spacers of extrachromosomal origin. *Microbiology*, 151(8):2551–2561, 2005.

L. Bonstein, S. Elias, and D. Frank. Paraxial-fated mesoderm is required for neural crest induction in xenopus embryos. *Developmental Biology*, 193(2):156 – 168, 1998.

P. Bork and E. V. Koonin. Predicting functions from protein sequences—where are the bottlenecks? *Nature Genetics*, 18(4):313–318, Apr. 1998.

E. A. Bosman, A. C. Penn, J. C. Ambrose, R. Kettleborough, D. L. Stemple, and K. P. Steel. Multiple mutations in mouse Chd7 provide models for CHARGE syndrome. *Human Molecular Genetics*, 14(22):3463–3476, Nov. 2005.

A. Brehm, K. R. Tufteland, R. Aasland, and P. B. Becker. The many colours of chromodomains. *BioEssays*, 26(2):133–140, Feb. 2004.

M. E. Bronner and N. M. LeDouarin. Development and evolution of the neural crest: An overview. *Developmental Biology*, 366(1):2 – 9, 2012. ISSN 0012-1606.

M. Bronner-Fraser and S. Fraser. Developmental potential of avian trunk neural crest cells in situ. *Neuron*, 3(6):755–766, Dec. 1989.

M. Bronner-Fraser and S. E. Fraser. Cell lineage analysis reveals multipotency of some avian neural crest cells. *Nature*, 335(6186):161–164, Sept. 1988.

M. Bronner-Fraser and S. E. Fraser. Cell lineage analysis of the avian neural crest. *Development*, Suppl 2:17–22, 1991.

W. R. A. Brown, S. J. Hubbard, C. Tickle, and S. A. Wilson. The chicken as a model for large-scale analysis of vertebrate gene function. *Nature Reviews Genetics*, 4(2): 87–98, Feb. 2003.

J. D. Buenrostro, P. G. Giresi, L. C. Zaba, H. Y. Chang, and W. J. Greenleaf. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature Methods*, 10 (12):1213–1218, Oct. 2013.

E. Buitrago-Delgado, K. Nordin, A. Rao, L. Geary, and C. LaBonne. Shared regulatory programs suggest retention of blastula-stage potential in neural crest cells. *Science*, 348(6241):1332–1335, June 2015.

D. W. Burt. Chicken genome: current status and future opportunities. *Genome Research*, 15(12):1692–1698, Dec. 2005.

E. Calo and J. Wysocka. Modification of enhancer chromatin: What, how, and why? *Molecular Cell*, 49(5):1097–2765, 2018.

P. Cartwright, C. McLean, A. Sheppard, D. Rivett, K. Jones, and S. Dalton. LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. *Development*, 132(5):885–896, Mar. 2005.

T. Cermak, E. L. Doyle, M. Christian, L. Wang, Y. Zhang, C. Schmidt, J. A. Baller, N. V. Somia, A. J. Bogdanove, and D. F. Voytas. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Research*, 39(12):e82, July 2011.

K. K. Chan, Y. S. Chen, T. O. Yau, M. Fu, V. C. H. Lui, P. K. H. Tam, and M. H. Sham. Hoxb3 vagal neural crest-specific enhancer element for controlling enteric nervous system development. *Developmental Dynamics*, 233(2):473–483, June 2005.

C.-J. Chang, C.-H. Chao, W. Xia, J.-Y. Yang, Y. Xiong, C.-W. Li, W.-H. Yu, S. K. Rehman, J. L. Hsu, H.-H. Lee, M. Liu, C.-T. Chen, D. Yu, and M.-C. Hung. p53 regulates epithelial-mesenchymal transition and stem cell properties through modulating miRNAs. *Nature Cell Biology*, 13(3):317–323, Mar. 2011.

T. Chen and S. Y. R. Dent. Chromatin modifiers and remodellers: regulators of cellular differentiation. *Nature Reviews Genetics*, 15(2):93–106, Feb. 2014.

H. M. T. Choi, C. R. Calvert, N. Husain, D. Huss, J. C. Barsi, B. E. Deverman, R. C. Hunter, M. Kato, S. M. Lee, A. C. T. Abelin, A. Z. Rosenthal, O. S. Akbari, Y. Li, B. A. Hay, P. W. Sternberg, P. H. Patterson, E. H. Davidson, S. K. Mazmanian, D. A. Prober, M. van de Rijn, J. R. Leadbetter, D. K. Newman, C. Readhead, M. E. Bronner, B. Wold, R. Lansford, T. Sauka-Spengler, S. E. Fraser, and N. A. Pierce. Mapping a multiplexed zoo of mRNA expression. *Development*, 143(19):3632–3637, Oct. 2016.

E. Choi-Rhee and J. E. Cronan. The Biotin Carboxylase-Biotin Carboxyl Carrier Protein Complex of Escherichia coliAcetyl-CoA Carboxylase. *Journal of Biological Chemistry*, 278(33):30806–30812, Aug. 2003.

C. R. Clapier and B. R. Cairns. The Biology of Chromatin Remodeling Complexes. *Annual Review of Biochemistry*, 78(1):273–304, June 2009.

E. B. Clark. Pathogenetic mechanisms of congenital cardiovascular malformations revisited. *Seminars in Perinatology*, 20(6):465–472, Dec. 1996.

L. Cong, F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. D. Hsu, X. Wu, W. Jiang, L. A. Marraffini, and F. Zhang. Multiplex genome engineering using crispr/cas systems. *Science*, 339(6121):819–823, 2013.

S. J. Conway, D. J. Henderson, and A. J. Copp. Pax3 is required for cardiac neural crest migration in the mouse: evidence from the splotch (Sp2H) mutant. *Development*, 124(2):505–514, Jan. 1997.

D. R. Corey and J. M. Abrams. Morpholino antisense oligonucleotides: tools for investigating vertebrate development. *Genome Biology*, 2(5):Reviews 1015.1–1015.3, 2001.

T. L. Creazzo, R. E. Godt, L. Leatherbury, S. J. Conway, and M. L. Kirby. Role of cardiac neural crest cells in cardiovascular development. *Annual Review of Physiology*, 60(1):267–286, 1998.

M. P. Creyghton, A. W. Cheng, G. G. Welstead, T. Kooistra, B. W. Carey, E. J. Steine, J. Hanna, M. A. Lodato, G. M. Frampton, P. A. Sharp, L. A. Boyer, R. A. Young, and R. Jaenisch. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *PNAS*, 107(50):21931–21936, Dec. 2010.

E. H. Davidson and M. S. Levine. Gene Regulatory Networks. *PNAS*, pages 1–1, Mar. 2005.

E. H. Davidson and M. S. Levine. Properties of developmental gene regulatory networks. *PNAS*, pages 1–4, Dec. 2008.

E. H. Davidson, D. R. McClay, and L. Hood. Regulatory gene networks and the properties of the developmental process. *PNAS*, 100(4):1475–1480, Feb. 2003.

J. O. J. Davies, J. M. Telenius, S. J. McGowan, N. A. Roberts, S. Taylor, D. R. Higgs, and J. R. Hughes. Multiplexed analysis of chromosome conformation at vastly improved sensitivity. *Nature Methods*, 13(1):74–80, Jan. 2016.

E. de Boer, P. Rodriguez, E. Bonte, J. Krijgsveld, E. Katsantoni, A. Heck, F. Grosveld, and J. Strouboulis. Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. *PNAS*, 100(13): 7480–7485, June 2003.

Deal and Henikoff. A Simple Method for Gene Expression and Chromatin Profiling of Individual Cell Types within a Tissue. *Developmental Cell*, 21(6):1–1, 2011.

E. Deltcheva, K. Chylinski, C. M. Sharma, K. Gonzales, Y. Chao, Z. A. Pirzada, M. R. Eckert, J. Vogel, and E. Charpentier. Crispr rna maturation by trans-encoded small rna and host factor rnase iii. *Nature*, 471:602–607, 2011.

D. Diaczok, S. DiVall, I. Matsuo, F. E. Wondisford, A. M. Wolfe, and S. Radovick. Deletion of Otx2in GnRH Neurons Results in a Mouse Model of Hypogonadotropic Hypogonadism. *Molecular Endocrinology*, 25(5):833–846, May 2011.

R. M. Dirks and N. A. Pierce. Triggered amplification by hybridization chain reaction. *PNAS*, 101(43):15275–15278, Oct. 2004.

J. R. Dixon, I. Jung, S. Selvaraj, Y. Shen, J. E. Antosiewicz-Bourget, A. Y. Lee, Z. Ye, A. Kim, N. Rajagopal, W. Xie, Y. Diao, J. Liang, H. Zhao, V. V. Lobanenkov, J. R. Ecker, J. A. Thomson, and B. Ren. Chromatin architecture reorganization during stem cell differentiation. *Nature*, 518(7539):331–336, Feb. 2015.

M. J. Dixon, M. L. Marazita, T. H. Beaty, and J. C. Murray. Cleft lip and palate: understanding genetic and environmental influences. *Nature Reviews Genetics*, 12 (3):167–178, Mar. 2011.

M. L. Donnelly, L. E. Hughes, G. Luke, H. Mendoza, E. ten Dam, D. Gani, and M. D. Ryan. The 'cleavage' activities of foot-and-mouth disease virus 2A site-directed

mutants and naturally occurring '2A-like' sequences. *Journal of General Virology*, 82(Pt 5):1027–1041, May 2001.

N. Douarin and C. Kalcheim. *The Neural Crest*. Developmental and Cell Biology Series. Cambridge University Press, 1999. ISBN 9780521620109.

E. Dupin and L. Sommer. Neural crest progenitors and stem cells: From early development to adulthood. *Developmental Biology*, 366(1):83–95, June 2012.

E. Dupin, S. Creuzet, and N. M. Le Douarin. *The Contribution of the Neural Crest to the Vertebrate Body*, pages 96–119. Springer US, Boston, MA, 2006. ISBN 978-0-387-46954-6.

E. S. Emison, A. S. McCallion, C. S. Kashuk, R. T. Bush, E. Grice, S. Lin, M. E. Portnoy, D. J. Cutler, E. D. Green, and A. Chakravarti. A common sex-dependent mutation in a RET enhancer underlies Hirschsprung disease risk. *Nature*, 434(7035):857–863, Apr. 2005.

F. Emmert-Streib, M. Dehmer, and B. Haibe-Kains. Gene regulatory networks and their applications: understanding biological and medical problems in terms of networks. *Frontiers in cell and developmental biology*, 2(281):38, 2014a.

F. Emmert-Streib, M. Dehmer, and B. Haibe-Kains. Gene regulatory networks and their applications: understanding biological and medical problems in terms of networks. *Frontiers in Cell and Developmental Biology*, pages 1–7, Aug. 2014b.

E. Engelen, U. Akinci, J. C. Bryne, J. Hou, C. Gontan, M. Moen, D. Szumska, C. Kockx, W. van Ijcken, D. H. W. Dekkers, J. Demmers, E.-J. Rijkers, S. Bhattacharya, S. Philipsen, L. H. Pevny, F. G. Grosveld, R. J. Rottier, B. Lenhard, and R. A. Poot. Sox2 cooperates with Chd7 to regulate genes that are mutated in human syndromes. *Nature Genetics*, 43(6):607–611, June 2011.

D. J. Epstein. *Cis*-regulatory mutations in human disease. *Briefings in Functional Genomics and Proteomics*, 8(4):310–316, Sept. 2009.

J. Fantes, N. K. Ragge, S.-A. Lynch, N. I. McGill, J. R. O. Collin, P. N. Howard-Peebles, C. Hayward, A. J. Vivian, K. Williamson, V. van Heyningen, and D. R. Fitzpatrick. Mutations in SOX2 cause anophthalmia. *Nature Genetics*, 33(4):461–463, Apr. 2003.

W. Feng, D. Kawauchi, H. Körkel-Qu, H. Deng, E. Serger, L. Sieber, J. Lieberman, S. Jimeno-González, S. Lambo, B. Hanna, Y. Harim, M. Jansen, A. Neuerburg, O. Friesen, M. Zuckermann, V. Rajendran, J. Gronych, O. Ayrault, A. Korshunov, D. Jones, M. Kool, P. Northcott, P. Lichter, F. Cortés-Ledesma, S. Pfister, and H.-K. Liuc. Chd7 is indispensable for mammalian brain development through activation of a neuronal differentiation programme. *Nature Communications*, 8:14758, 2017.

W. Feng, C. Shao, and H.-K. Liu. Versatile roles of the chromatin remodeller chd7 during brain development and disease. *Frontiers in Molecular Neuroscience*, 10:309, 2017.

E. Frank and J. R. Sanes. Lineage of neurons and glia in chick dorsal root ganglia: analysis *in vivo* with a recombinant retrovirus. *Development*, 111(4):895–908, Apr. 1991.

K. Fukada. Hormonal control of neurotransmitter choice in sympathetic neurone cultures. *Nature*, 287(5782):553–555, Oct. 1980.

E. E. M. Furlong and M. Levine. Developmental enhancers and chromosome topology. *Science*, 361(6409):1341–1345, 2018.

L. S. Gammill. Genomic analysis of neural crest induction. *Development*, 129(24): 5731–5741, Dec. 2002.

M. I. García-Castro, C. Marcelle, and M. Bronner-Fraser. Ectodermal wnt function as a neural crest inducer. *Science*, 297(5582):848–851, 2002.

G. K. Geiss, R. E. Bumgarner, B. Birditt, T. Dahl, N. Dowidar, D. L. Dunaway, H. P. Fell, S. Ferree, R. D. George, T. Grogan, J. J. James, M. Maysuria, J. D. Mitton, P. Oliveri, J. L. Osborn, T. Peng, A. L. Ratcliffe, P. J. Webster, E. H. Davidson, L. Hood, and K. Dimitrov. Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nature Biotechnology*, 26(3):317–325, Mar. 2008.

M. G. Goll and T. H. Bestor. Eukaryotic Cytosine Methyltransferases. *Annual Review of Biochemistry*, 74(1):481–514, June 2005.

M. L. Golson and K. H. Kaestner. Fox transcription factors: from development to disease. *Development*, 143(24):4558–4570, Dec. 2016.

C. T. Gordon, S. Benko, J. Amiel, and S. Lyonnet. *Cis*-regulatory disruption at the SOX9 locus as a cause of Pierre Robin sequence. In *Gene Regulatory Sequences and Human Disease*, pages 123–136. Springer New York, New York, NY, Sept. 2011.

A. Graham and S. M. Shimeld. The origin and evolution of the ectodermal placodes. *Journal of Anatomy*, 222(1):32–40, Jan. 2013.

A. C. Groner, S. Meylan, A. Ciuffi, N. Zangger, G. Ambrosini, N. Dénervaud, P. Bucher, and D. Trono. KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. *PLoS Genetics*, 6(3):e1000869, Mar. 2010.

Y. Guo and W. Qiang. CTCF/cohesin-mediated DNA looping is required for protocadherin alpha promoter choice. *PNAS*, pages 1–15, Nov. 2015.

Y. Guo, Q. Xu, D. Canzio, J. Shou, J. Li, D. U. Gorkin, I. Jung, H. Wu, Y. Zhai, Y. Tang, Y. Lu, Y. Wu, Z. Jia, W. Li, M. Q. Zhang, B. Ren, A. R. Krainer, T. Maniatis, and Q. Wu. CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell*, 162(4):900–910, Aug. 2015.

B. K. Hall. The neural crest as a fourth germ layer and vertebrates as quadroblastic not triploblastic. *Evolution & Development*, 2(1):3–5, 2000.

V. Hamburger and H. L. Hamilton. A series of normal stages in the development of the chick embryo. *Journal of Morphology*, 1951.

D. C. Hargreaves and G. R. Crabtree. ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. *Cell Research*, pages 1–25, Mar. 2011.

D. Hay, J. R. Hughes, C. Babbs, J. O. Davies, B. J. Graham, L. Hanssen, M. T. Kassouf, A. M. M. Oudelaar, J. A. Sharpe, M. C. Suciu, J. Telenius, R. Williams, C. Rode, P.-S. Li, L. A. Pennacchio, J. A. Sloane-Stanley, H. Ayyub, S. Butler, T. Sauka-Spengler, R. J. Gibbons, A. J. Smith, W. G. Wood, and D. R. Higgs. Genetic dissection of the a-globin super-enhancer *in vivo*. *Nature Genetics*, 48(8):895–903, 2016.

A. He and W. T. Pu. *Genome-Wide Location Analysis by Pull Down of In Vivo Biotinylated Transcription Factors*. John Wiley & Sons, Inc., Hoboken, NJ, USA, May 2001.

N. D. Heintzman, R. K. Stuart, G. Hon, Y. Fu, C. W. Ching, R. D. Hawkins, L. O. Barrera, S. V. Calcar, C. Qu, K. A. Ching, W. Wang, Z. Weng, R. D. Green, G. E. Crawford, and B. Ren. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature Genetics*, 39:311–318, 2007.

B. Hendrich and W. Bickmore. Human diseases with underlying defects in chromatin structure and modification. *Human Molecular Genetics*, 10(20):2233–2242, Oct. 2001.

P. D. Henion and J. A. Weston. Timing and pattern of cell fate restrictions in the neural crest lineage. *Development*, 124(21):4351–4359, Nov. 1997.

E. E. Henninger and Z. F. Pursell. DNA polymerase E and its roles in genome stability. *IUBMB Life*, 66(5):339–351, May 2014.

W. His. *Untersuchungen über die erste Anlage des Wirbelthierleibes die erste Entwickelung des Hühnchens im Ei* . F.C.W. Vogel,, Leipzig :, 1868.

L. Ho and G. R. Crabtree. Chromatin remodelling during development. *Nature*, 463 (7280):474–484, 2017.

A. Holmberg, A. Blomstergren, O. Nord, M. Lukacs, J. Lundeberg, and M. Uhlén. The biotin-streptavidin interaction can be reversibly broken using water at elevated temperatures. *Electrophoresis*, 26(3):501–510, Feb. 2005.

P. Hsu, A. Ma, M. Wilson, G. Williams, J. Curotta, C. F. Munns, and S. Mehr. Charge syndrome: A review. *Journal of Paediatrics and Child Health*, 50(7):504–511, 2014a.

P. D. Hsu, E. S. Lander, and F. Zhang. Development and applications of crispr-cas9 for genome engineering. *Developmental Dynamics*, 157(6):1262–1278, 2014b.

N. Hu, P. Strobl-Mazzulla, T. Sauka-Spengler, and M. E. Bronner. DNA methyltransferase3A as a molecular switch mediating the neural tube-to-neural crest fate transition. *Genes and Development*, 26(21):2380–2385, Nov. 2012.

C. Huang, M.-C. Kratzer, D. Wedlich, and J. Kashef. E-cadherin is required for cranial neural crest migration in Xenopus laevis. *Developmental Biology*, 411(2):159–171, Mar. 2016.

J. R. Hughes, N. Roberts, S. McGowan, D. Hay, E. Giannoulatou, M. Lynch, M. D. Gobbi, S. Taylor, R. Gibbons, and D. R. Higgs. Analysis of hundreds of *cis*-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nature Genetics*, 46:205–212, 2014.

E. A. Hurd, P. L. Capers, M. N. Blauwkamp, M. E. Adams, Y. Raphael, H. K. Poucher, and D. M. Martin. Loss of Chd7 function in gene-trapped reporter mice is embryonic lethal and associated with severe defects in multiple developing tissues. *Mammalian Genome*, 18(2):94–104, Feb. 2007.

N. Itasaki, S. Bel-Vialar, and R. Krumlauf. 'Shocking' developments in chick embryology: electroporation and in ovo gene expression. *Nature Cell Biology*, 1:1–5, Nov. 1999.

N. Janssen, J. E. H. Bergman, M. A. Swertz, L. Tranebjaerg, M. Lodahl, J. Schoots, R. M. W. Hofstra, C. M. A. van Ravenswaaij-Arts, and L. H. Hoefsloot. Mutation

update on the CHD7 gene involved in CHARGE syndrome. *Human Mutation*, 33 (8):1149–1160, Aug. 2012.

M. Jeanpierre, C. Turleau, A. Aurias, M. Prieur, F. Ledeist, A. Fischer, and E. Viegas-Pequignot. An embryonic-like methylation pattern of classical satellite DNA is observed in ICF syndrome. *Human Molecular Genetics*, 2(6):731–735, June 1993.

J. C. Jemc. Somatic gonadal cells: The supporting cast for the germline. *Genesis*, 49 (10):753–775, 2011.

M. Jiang, J. Stanke, and J. M. Lahti. The connections between neural crest development and neuroblastoma. *Current Topics in Developmental Biology*, page 77–127, 2011.

M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, and E. Charpentier. A programmable dual-rna–guided dna endonuclease in adaptive bacterial immunity. *Science*, 337(6096):816–821, 2012. doi: 10.1126/science.1225829.

M. Jongmans, C. Van Ravenswaaij-Arts, N. Pitteloud, T. Ogata, N. Sato, H. Claahsen-van der Grinten, K. Van Der Donk, S. Seminara, J. Bergman, H. Brunner, W. Crowley, and L. Hoefsloot. Chd7 mutations in patients initially diagnosed with kallmann syndrome – the clinical overlap with charge syndrome. *Clinical Genetics*, 75(1):65–71, 2008.

M. C. J. Jongmans, R. J. Admiraal, K. P. van der Donk, L. E. L. M. Vissers, A. F. Baas, L. Kapusta, J. M. van Hagen, D. Donnai, T. J. de Ravel, J. A. Veltman, A. Geurts van Kessel, B. B. A. De Vries, H. G. Brunner, L. H. Hoefsloot, and C. M. A. van Ravenswaaij. CHARGE syndrome: the phenotypic spectrum of mutations in the CHD7 gene. *Journal of Medical Genetics*, 43(4):306–314, Apr. 2006.

C. Kalcheim and D. Sela-Donenfeld. Inhibition of noggin expression in the dorsal neural tube by somitogenesis: a mechanism for coordinating the timing of neural crest emigration. *Development*, pages 1–10, Oct. 2000.

R. N. Kelsh and C. A. Erickson. *Neural Crest: Origin, Migration and Differentiation*, volume 139. John Wiley & Sons, Ltd, Chichester, UK, 2 edition, May 2001.

A. Keyte and M. R. Hutson. The neural crest in cardiac congenital anomalies. *Differentiation*, 84(1):25–40, July 2012.

M. A. F. Khan, L. M. Soto-Jimenez, T. Howe, A. Streit, A. Sosinsky, and C. D. Stern. Computational tools and resources for prediction and analysis of gene regulatory regions in the chick genome. *Genesis*, 51(5):311–324, Mar. 2013.

J. Khudyakov and M. Bronner-Fraser. Comprehensive spatiotemporal analysis of early chick neural crest network genes. *Developmental Dynamics*, 238(3):716–723, Mar. 2009.

H.-G. Kim, I. Kurth, F. Lan, I. Meliciani, W. Wenzel, S. H. Eom, G. B. Kang, G. Rosenberger, M. Tekin, M. Ozata, D. P. Bick, R. J. Sherins, S. L. Walker, Y. Shi, J. F. Gusella, and L. C. Layman. Mutations in chd7, encoding a chromatin-remodeling protein, cause idiopathic hypogonadotropic hypogonadism and kallmann syndrome. *The American Journal of Human Genetics*, 83(4):511 – 519, 2008.

J. Kim, A. B. Cantor, S. H. Orkin, and J. Wang. Use of *in vivo* biotinylation to study protein–protein and protein–DNA interactions in mouse embryonic stem cells. *Nature Protocols*, 4(4):506–517, Mar. 2009.

M. L. Kirby. Cardiac morphogenesis–recent research advances. *Pediatric Research*, 21 (3):219–224, Mar. 1987.

M. L. Kirby and K. L. Waldo. Role of neural crest in congenital heart disease. *Circulation*, 82(2):332–340, Aug. 1990.

M. L. Kirby, D. C. Conrad, and D. E. Stewart. Increase in the cholinergic cardiac plexus in sympathetically aneural chick hearts. *Cell and Tissue Research*, 247(3): 489–496, Mar. 1987.

K. E. Kolodziej, F. Pourfarzad, E. de Boer, S. Krpic, F. Grosveld, and J. Strouboulis. Optimal use of tandem biotin and V5 tags in ChIP assays. *BMC Molecular Biology*, 10(1):6–14, 2009.

P. Kolovos, T. A. Knoch, F. G. Grosveld, P. R. Cook, and A. Papantonis. Enhancers and silencers: an integrated and simple model for their function. *Epigenetics & chromatin*, 5(1):1, Jan. 2012.

R. Kos, M. V. Reedy, R. L. Johnson, and C. A. erickson. The winged-helix transcription factor FoxD3 is important for establishing the neural crest lineage and repressing melanogenesis in avian embryos. *Development*, 128(8):1467–1479, Apr. 2001.

T. Kudo and S. Sutou. Usage of putative chicken U6 promoters for vector-based RNA interference. *Journal of Reproduction and Development*, 51(3):411–417, June 2005.

P. M. Kulesa, J. A. Morrison, and C. M. Bailey. The neural crest and cancer: A developmental spin on melanoma. *Cells, Tissues, Organs*, pages 1422–6405, 2013.

E. Z. Kvon. Induction by medial and lateral pieces of archenteron roof, with special reference to the determination of the neural crest. *Acta Neerl. Norm. Pathol.*, 55: 348–362, 1945.

E. Z. Kvon. Using transgenic reporter assays to functionally characterize enhancers in animals. *Genomics*, 106(3):185–192, Sept. 2015.

D. Z. Kwaku Dad Abu-Bonsrah and D. F. Newgreen. Crispr/cas9 targets chicken embryonic somatic cells *In Vitro* and *In Vivo* and generates phenotypic abnormalities. *Scientific Reports*, 6:34524, 2016.

C. LaBonne and M. Bronner-Fraser. Neural crest induction in Xenopus: evidence for a two-signal model. *Development*, 125(13):2403, July 1998.

R. Lahav, C. Ziller, E. Dupin, and N. M. Le Douarin. Endothelin 3 promotes neural crest cell proliferation and mediates a vast increase in melanocyte number in culture. *PNAS*, 93(9):3892–3897, Apr. 1996.

S. R. Lalani, A. M. Safiullah, S. D. Fernbach, K. G. Harutyunyan, C. Thaller, L. E. Peterson, J. D. McPherson, R. A. Gibbs, L. D. White, M. Hefner, S. L. H. Davenport, J. M. Graham, C. A. Bacino, N. L. Glass, J. A. Towbin, W. J. Craigen, S. R. Neish, A. E. Lin, and J. W. Belmont. Spectrum of CHD7 mutations in 110 individuals with CHARGE syndrome and genotype-phenotype correlation. *American Journal of Human Genetics*, 78(2):303–314, Feb. 2006.

G. Langst and L. Manelyte. Chromatin Remodelers: From Function to Dysfunction. *Genes*, 6(2):299–324, June 2015.

W. S. Layman, E. A. Hurd, and D. M. Martin. Chromodomain proteins in development: lessons from CHARGE syndrome. *Clinical Genetics*, 78(1):11–20, Apr. 2010.

N. Le Douarin. A biological cell labeling technique and its use in experimental embryology. *Developmental Biology*, 30(1):217–222, 1973.

N. Le Douarin and C. Kalcheim. *The Neural Crest*. Cambridge University Press, 1999.

N. M. Le Douarin. The avian embryo as a model to study the development of the neural crest: a long and still ongoing story. *Mechanisms of Development*, 121(9): 1089–1102, Sept. 2004.

N. M. Le Douarin and E. Dupin. Multipotentiality of the neural crest. *Current Opinion in Genetics & Development*, 13(5):529–536, Oct. 2003.

N. M. Le Douarin and M.-A. Teillet. The migration of neural crest cells to the wall of the digestive tract in avian embryo. *Development*, 30(1):31–48, Aug. 1973.

C. S. Le Lievre and N. M. Le Douarin. Mesenchymal derivatives of the neural crest: analysis of chimaeric quail and chick embryos. *Journal of Embryology and Experimental Morphology*, 34(1):125–154, Aug. 1975.

O. J. Lehmann, J. C. Sowden, P. Carlsson, T. Jordan, and S. S. Bhattacharya. Fox's in development and disease. *Trends in Genetics*, 19(6):339–344, June 2003.

K. F. Liem, G. Tremml, H. Roelink, and T. M. Jessell. Dorsal differentiation of neural plate cells induced by bmp-mediated signals from epidermal ectoderm. *Cell*, 82 (6):969 – 979, 1995.

A. E. Lin, J. R. Siebert, and J. M. Graham. Central nervous system malformations in the charge association. *American Journal of Medical Genetics*, 37(3):304–310, 1990.

Y. Lin, X.-Y. Li, A. L. Willis, C. Liu, G. Chen, and S. J. Weiss. Snail1-dependent control of embryonic stem cell pluripotency and lineage commitment. *Nature Communications*, 5:3070, 2014.

Y. Liu and P. A. Labosky. Regulation of embryonic stem cell self-renewal and pluripotency by Foxd3. *Stem Cells*, 26(10):2475–2484, Oct. 2008.

Y. Liu and A. Xiao. Epigenetic regulation in neural crest development. *Birth Defects Research Part A: Clinical and Molecular Teratology*, 91(8):788–796, May 2011.

M. Lukoseviciute, D. Gavriouchkina, R. M. Williams, T. Hochgreb-Hagele, U. Senanayake, V. Chong-Morrison, S. Thongjuea, E. Repapi, A. Mead, and T. Sauka-Spengler. From Pioneer to Repressor: Bimodal foxd3 Activity Dynamically Remodels Neural Crest Regulatory Landscape *In-Vivo*. *Developmental cell*, 47 (5):608–628.e6, Dec. 2018.

A. Lumsden, N. Sprawson, and A. Graham. Segmental origin and migration of neural crest cells in the hindbrain region of the chick embryo. *Development*, 113 (4):1281–1291, Dec. 1991.

D. G. Lupiáñez, K. Kraft, V. Heinrich, P. Krawitz, F. Brancati, E. Klopocki, D. Horn, H. Kayserili, J. M. Opitz, R. Laxova, F. Santos-Simarro, B. Gilbert-Dussardier, L. Wittler, M. Borschiwer, S. A. Haas, M. Osterwalder, M. Franke, B. Timmermann, J. Hecht, M. Spielmann, A. Visel, and S. Mundlos. Disruptions of Topological

Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions. *Cell*, 161(5):1012–1025, May 2015.

L. H. Maguire, A. R. Thomas, and A. M. Goldstein. Tumors of the neural crest: Common themes in development and cancer. *Developmental Dynamics*, 244(3): 311–322, 2015.

P. Mali, L. Yang, K. M. Esvelt, J. Aach, M. Guell, J. E. DiCarlo, J. E. Norville, and G. M. Church. Rna-guided human genome engineering via cas9. *Science*, 339 (6121):823–826, 2013.

S. A. Mani, W. Guo, M.-J. Liao, E. N. Eaton, A. Ayyanan, A. Y. Zhou, M. Brooks, F. Reinhard, C. C. Zhang, M. Shipitsin, L. L. Campbell, K. Polyak, C. Brisken, J. Yang, and R. A. Weinberg. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell*, 133(4):704–715, May 2008.

P. Maraschio, O. Zuffardi, T. D. Fior, and L. Tiepolo. Immunodeficiency, centromeric heterochromatin instability of chromosomes 1, 9, and 16, and facial anomalies: the ICF syndrome. *Journal of Medical Genetics*, 25(3):173–180, Mar. 1988.

L. Marchant, C. Linker, P. Ruiz, N. Guerrero, and R. Mayor. The inductive properties of mesoderm suggest that the neural crest cells are specified by a bmp gradient. *Developmental Biology*, 198(2):319 – 329, 1998.

S. Marcos, J. Sarfati, C. Leroy, C. Fouveaut, P. Parent, C. Metz, S. Wolczynski, M. Gérard, E. Bieth, F. Kurtz, O. Verier-Mine, L. Perrin, F. Archambeaud, S. Cabrol, P. Rodien, H. Hove, T. Prescott, D. Lacombe, S. Christin-Maitre, P. Touraine, S. Hieronimus, D. Dewailly, J. Young, M. Pugeat, J.-P. Hardelin, and C. Dodé. The prevalence of chd7 missense versus truncating mutations is higher in patients with kallmann syndrome than in typical charge patients. *The Journal of Clinical Endocrinology & Metabolism*, 99(10):E2138–E2143, 2014.

K. Martins-Taylor, D. I. Schroeder, J. M. LaSalle, M. Lalande, and R.-H. Xu. Role of DNMT3B in the regulation of early neural and neural crest specifiers. *Epigenetics*, 7(1):71–82, Jan. 2012.

M. Matsumata, M. Uchikawa, Y. Kamachi, and H. Kondoh. Multiple N-cadherin enhancers identified by systematic functional screening indicate its Group B1 SOX-dependent regulation in neural and placodal development. *Developmental Biology*, 286(2):601–617, Oct. 2005.

L. A. J. M. Matthias Dehmer and F. Emmert-Streib. Quantitative network measures as biomarkers for classifying prostate cancer disease states: a systems approach to diagnostic biomarkers. *PLoS ONE*, 8(11):e77602, 2013.

L. Menendez, M. J. Kulik, A. T. Page, S. S. Park, J. D. Lauderdale, M. L. Cunningham, and S. Dalton. Directed differentiation of human pluripotent cells to neural crest stem cells. *Nature Protocols*, 8(1):203–212, Jan. 2013.

D. Meulemans and M. Bronner-Fraser. Gene-regulatory interactions in neural crest evolution and development. *Developmental Cell*, 7(3):291–299, Sept. 2004.

D. Meulemans and M. Bronner-Fraser. Central role of gene cooption in neural crest evolution. *Journal of Experimental Zoology, Part B: Molecular and Developmental Evolution*, 304(4):298–303, July 2005.

C. Milet and A. H. Monsoro-Burq. Neural crest induction at the neural plate border in vertebrates. *Developmental Biology*, 366(1):22 – 33, 2012.

K. Møllgård, A. Jespersen, M. C. Lutterodt, C. Yding Andersen, P. E. Høyer, and A. G. Byskov. Human primordial germ cells migrate along nerve fibers and Schwann cells from the dorsal hind gut mesentery to the gonadal ridge. *MHR: Basic science of reproductive medicine*, 16(9):621–631, Aug. 2010.

A.-H. Monsoro-Burq, R. B. Fletcher, and R. M. Harland. Neural crest induction by paraxial mesoderm in Xenopus embryos requires FGF signals. *Development*, 130 (>14):3111–3124, 2003.

A.-P. Morel, M. Lievre, C. Thomas, G. Hinkal, S. Ansieau, and A. Puisieux. Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PLoS ONE*, 3(8):e2888, Aug. 2008.

J. D. Moury and A. G. Jacobson. Neural fold formation at newly created boundaries between neural plate and epidermis in the axolotl. *Developmental Biology*, 133(1): 44–57, 1989.

N. A. Mundell and P. A. Labosky. Neural crest stem cell multipotency requires Foxd3 to maintain neural potential and repress mesenchymal fates. *Development*, 138(4):641–652, Jan. 2011.

T. Muramatsu, Y. Mizutani, and J.-i. Okumura. Live Detection of the Firefly Luciferase Gene Expression Chicken by Bioluminescence in Incubating Embryos . *Animal Science Technology Japan*, pages 1–4, Dec. 1996.

T. Muramatsu, Y. Mizutani, Y. Ohmori, and J. Okumura. Comparison of three nonviral transfection methods for foreign gene expression in early chicken embryos in ovo. *Biochemical and biophysical research communications*, 230(2):376–380, Jan. 1997.

T. Muramatsu, T. Imai, H.-M. Park, H. Watanabe, A. Nakamura, and J.-i. Okumura. Gene gun-mediated *in vivo* analysis of tissue- specific repression of gene transcription driven by the chicken ovalbumin promoter in the liver and oviduct of laying hens . *Molecular and Cellular Biochemistry*, pages 1–6, July 1998.

K. Musunuru. Genome editing of human pluripotent stem cells to generate human cellular disease models. *Disease Models & Mechanisms*, 6(4):896–904, July 2013.

R. Nakato and K. Shirahige. Recent advances in ChIP-seq analysis: from quality management to whole-genome annotation. *Briefings in Bioinformatics*, 18(2):279–290, Mar. 2017.

A. Narla and B. L. Ebert. Ribosomopathies: human disorders of ribosome dysfunction. *Blood*, 115(16):3196–3205, 2010. doi: 10.1182/blood-2009-10-178129.

G. J. Narlikar, R. Sundaramoorthy, and T. Owen-Hughes. Mechanisms and Functions of ATP-Dependent Chromatin-Remodeling Enzymes. *Cell*, 154(3):490–503, Aug. 2013.

P. Nestorov, H.-R. Hotz, Z. Liu, and A. H. F. M. Peters. Dynamic expression of chromatin modifiers during developmental transitions in mouse preimplantation embryos. *Scientific Reports*, 5:14347, Sept. 2015.

N. Nikitina, T. Sauka-Spengler, and M. Bronner-Fraser. Chapter 1 gene regulatory networks in neural crest development and evolution. In *Evolution and Development*, volume 86 of *Current Topics in Developmental Biology*, pages 1 – 14. Academic Press, 2009.

T. Nishi, K. Yoshizato, S. Yamashiro, H. Takeshima, K. Sato, K. Hamada, I. Kitamura, T. Yoshimura, H. Saya, J. Kuratsu, and Y. Ushio. High-efficiency *in vivo* gene transfer using intraarterial plasmid DNA injection following *in vivo* electroporation. *Cancer Research*, 56(5):1050–1055, Mar. 1996.

K. i. Nishijima and S. Iijima. Transgenic chickens. *Development, growth & differentiation*, 55(1):207–216, Dec. 2012.

E. Nitzan, S. Krispin, E. R. Pfaltzgraff, A. Klar, P. A. Labosky, and C. Kalcheim. A dynamic code of dorsal neural tube genes regulates the segregation between neurogenic and melanogenic neural crest cells. *Development*, 140(11):2269–2279, May 2013.

M. Okano, D. W. Bell, D. A. Haber, and E. Li. DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. *Cell*, 99(3):247–257, Oct. 1999.

C.-T. Ong and V. G. Corces. Enhancers: emerging roles in cell fate specification. *EMBO reports*, 13(5):423–430, May 2012.

C. P. Ordahl and N. M. Le Douarin. Two myogenic lineages within the developing somite. *Development*, 114(2):339–353, Feb. 1992.

J. Pachlopnik Schmid, R. Lemoine, N. Nehme, V. Cormier-Daire, P. Revy, F. Debeurme, M. Debré, P. Nitschke, C. Bole-Feysot, L. Legeai-Mallet, A. Lim, J.-P. de Villartay, C. Picard, A. Durandy, A. Fischer, and G. de Saint Basile. Polymerase e1 mutation in a human syndrome with facial dysmorphism, immunodeficiency, livedo, and short stature ("fils syndrome"). *Journal of Experimental Medicine*, 209 (13):2323–2330, 2012. doi: 10.1084/jem.20121303.

R. A. Pagon, J. M. Graham, J. Zonana, and S. L. Yong. Coloboma, congenital heart disease, and choanal atresia with multiple anomalies: CHARGE association. *The Journal of Pediatrics*, 99(2):223–227, Aug. 1981.

S. Pauli, R. Bajpai, and A. Borchers. CHARGEd with neural crest defects. *American Journal of Medical Genetics Part A*, 79:50–9, Oct. 2017.

H. D. Pomeranz, T. P. Rothman, and M. D. Gershon. Colonization of the post-umbilical bowel by cells derived from the sacral neural crest: direct tracing of cell migration using an intercalating probe and a replication-deficient retrovirus. *Development*, 111(3):647–655, Mar. 1991.

Z. F. Pursell and T. A. Kunkel. DNA Polymerase E: A Polymerase of Unusual Size (and Complexity). *Progress in Nucleic Acid Research and Molecular Biology*, pages 101–145, 2008.

F. Rahimov, M. L. Marazita, A. Visel, M. E. Cooper, M. J. Hitchler, M. Rubini, F. E. Domann, M. Govil, K. Christensen, C. Bille, M. Melbye, A. Jugessur, R. T. Lie, A. J. Wilcox, D. R. Fitzpatrick, E. D. Green, P. A. Mossey, J. Little, R. P. Steegers-Theunissen, L. A. Pennacchio, B. C. Schutte, and J. C. Murray. Disruption of an AP-2alpha binding site in an IRF6 enhancer is associated with cleft lip. *Nature Genetics*, 40(11):1341–1347, Oct. 2008.

D. W. Raible, A. Wood, W. Hodsdon, P. D. Henion, J. A. Weston, and J. S. Eisen. Segregation and early dispersal of neural crest cells in the embryonic zebrafish. *Developmental Dynamics*, 195(1):29–42, Sept. 1992.

T. Raivio. CHARGE syndrome and Kallmann syndrome: are the two genetically related? *Expert Review of Endocrinology & Metabolism*, 7(6):579–581, Jan. 2014.

V. Randall, K. McCue, C. Roberts, V. Kyriakopoulou, S. Beddow, A. N. Barrett, F. Vitelli, K. Prescott, C. Shaw-Smith, K. Devriendt, E. Bosman, G. Steffes, K. P. Steel, S. Simrick, M. A. Basson, E. Illingworth, and P. J. Scambler. Great vessel development requires biallelic expression of chd7 and tbx1 in pharyngeal ectoderm in mice. *The Journal of Clinical Investigation*, 119(11):3301–3310, 11 2009.

M. E. Rawles. Origin of melanophores and their role in development of color patterns in vertebrates. *Physiological Reviews*, 28(4):383–408, Oct. 1948.

K. R. Rosenbloom, C. A. Sloan, V. S. Malladi, T. R. Dreszer, K. Learned, V. M. Kirkup, M. C. Wong, M. Maddren, R. Fang, S. G. Heitner, B. T. Lee, G. P. Barber, R. A. Harte, M. Diekhans, J. C. Long, S. P. Wilder, A. S. Zweig, D. Karolchik, R. M. Kuhn, D. Haussler, and W. J. Kent. ENCODE data in the UCSC Genome Browser: year 5 update. *Nucleic Acids Research*, 41(Database issue):D56–63, Jan. 2013.

D. Rudra, P. deRoos, A. Chaudhry, R. E. Niec, A. Arvey, R. M. Samstein, C. Leslie, S. A. Shaffer, D. R. Goodlett, and A. Y. Rudensky. Transcription factor Foxp3 and its protein partners form a complex regulatory network. *Nature Immunology*, 13 (10):1010–1019, Aug. 2012.

A. Sandelin, W. Alkema, P. Engström, W. W. Wasserman, and B. Lenhard. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Research*, 32(Database issue):D91–4, Jan. 2004.

J. D. Sander and J. K. Joung. Crispr-cas systems for editing, regulating and targeting genomes. *Nature Biotechnology*, 32(4):347–355, 2014.

D. Sanlaville and A. Verloes. CHARGE syndrome: an update. *European Journal of Human Genetics : EJHG*, 15(4):389–399, Apr. 2007.

Sanlaville, D, Etchevers, H C, Gonzales, M, Martinovic, J, Clément-Ziza, M, Delezoide, A-L, Aubry, M-C, Pelet, A, Chemouny, S, Cruaud, C, Audollent, S, Esculpavit, C, Goudefroye, G, Ozilou, C, Fredouille, C, Joye, N, Morichon-Delvallez, N, Dumez, Y, Weissenbach, J, Munnich, A, Amiel, J, Encha-Razavi, F, Lyonnet, S, Vekemans, M, and Attié-Bitach, T. Phenotypic spectrum of CHARGE syndrome in fetuses with CHD7 truncating mutations correlates with expression during human development. *Journal of Medical Genetics*, 43(3):211–217, Mar. 2006.

N. Sasai, K. Mizuseki, and Y. Sasai. Requirement of FoxD3-class signaling for neural crest determination in Xenopus. *Development*, 128(13):2525–2536, July 2001.

S. Sato, K. Ikeda, G. Shioi, K. Nakao, H. Yajima, and K. Kawakami. Regulation of Six1 expression by evolutionarily conserved enhancers in tetrapods. *Developmental Biology*, 368(1):95–108, Aug. 2012.

T. Sauka-Spengler and M. Barembaum. Gain- and loss-of-function approaches in the chick embryo. *Methods in Cell Biology*, 87:237–256, 2008.

T. Sauka-Spengler and M. Bronner. Snapshot: neural crest. *Cell*, 143(3):486–486.e1, Oct. 2010.

T. Sauka-Spengler and M. Bronner-Fraser. A gene regulatory network orchestrates neural crest formation. *Nature Reviews Molecular Cell Biology*, 9(7):557–568, June 2008.

P. Savagner, K. M. Yamada, and J. P. Thiery. The zinc-finger protein slug causes desmosome dissociation, an initial and necessary step for growth factor-induced epithelial-mesenchymal transition. *The Journal of Cell Biology*, 137(6):1403–1419, June 1997.

P. J. Schatz. Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: a 13 residue consensus peptide specifies biotinylation in Escherichia coli. *Bio/technology (Nature Publishing Company)*, 11(10):1138–1143, Oct. 1993.

C. Scheel and R. A. Weinberg. Phenotypic plasticity and epithelial-mesenchymal transitions in cancer and normal stem cells. *International Journal of Cancer*, 129(10): 2310–2314, Nov. 2011.

G. Schlosser. Chapter four - making senses: Development of vertebrate cranial placodes. In K. Jeon, editor, *International Review of Cell and Molecular Biology*, volume 283 of *International Review of Cell and Molecular Biology*, pages 129 – 234. Academic Press, 2010.

M. P. Schnetz, C. F. Bartels, K. Shastri, D. Balasubramanian, G. E. Zentner, R. Balaji, X. Zhang, L. Song, Z. Wang, T. LaFramboise, G. E. Crawford, and P. C. Scacheri. Genomic distribution of CHD7 on chromatin tracks H3K4 methylation patterns. *Genome Research*, 19(4):590–601, Apr. 2009.

M. P. Schnetz, L. Handoko, B. Akhtar-Zaidi, C. F. Bartels, C. F. Pereira, A. G. Fisher, D. J. Adams, P. Flicek, G. E. Crawford, T. Laframboise, P. Tesar, C.-L. Wei, and P. C. Scacheri. CHD7 Targets Active Gene Enhancer Elements to Modulate ES Cell-Specific Gene Expression. *PLoS Genetics*, 6(7):e1001023–15, July 2010.

Y. Schulz, P. Wehner, L. Opitz, G. Salinas-Riester, E. M. H. F. Bongers, C. M. A. van Ravenswaaij-Arts, J. Wincent, J. Schoumans, J. Kohlhase, A. Borchers, and S. Pauli. CHD7, the gene mutated in CHARGE syndrome, regulates genes involved in neural crest cell guidance. *Human Genetics*, 133(8):997–1009, Aug. 2014.

M. Schwanzel-Fukuda, D. Bick, and D. W. Pfaff. Luteinizing hormone-releasing hormone (LHRH)-expressing cells do not migrate normally in an inherited hypogonadal (Kallmann) syndrome. *Molecular Brain Research*, 6(4):311–326, Dec. 1989.

M. A. Selleck and M. Bronner-Fraser. Origins of the avian neural crest: the role of neural plate-epidermal interactions. *Development*, 121(2):525, Feb. 1995.

T. Sengoku, O. Nureki, A. Nakamura, S. Kobayashi, and S. Yokoyama. Structural basis for RNA unwinding by the DEAD-box protein Drosophila Vasa. *Cell*, 125(2): 287–300, Apr. 2006.

H. W. Seo, T. M. Kim, J. W. Choi, B. K. Han, G. Song, and J. Y. Han. Evaluation of combinatorial cis-regulatory elements for stable gene expression in chicken cells. *BMC Biotechnology*, 10(1):69–8, 2010.

D. A. Shafritz. Use of Electroporation To Introduce Biologically Active Foreign Genes into Primary Rat Hepatocytes . *Molecular and Cellular Biology*, pages 1–3, Feb. 1986.

P. Sharma, F. Yan, V. A. Doronina, H. Escuin-Ordinas, M. D. Ryan, and J. D. Brown. 2A peptides provide distinct solutions to driving stop-carry on translational recoding. *Nucleic Acids Research*, 40(7):3143–3151, Apr. 2012.

F. M. V. Shashank Gandhi, Michael L Piacentino and M. E. Bronner. Optimization of crispr/cas9 genome editing for loss-of-function in the early chick embryo. *Developmental Biology*, pages 86–97, 2016.

D. Shlyueva, G. Stampfel, and A. Stark. Transcriptional enhancers: from properties to genome-wide predictions. *Nature Reviews Genetics*, 15(4):272–286, Apr. 2014.

J. R. Siebert, J. M. Graham, and C. MacDonald. Pathologic Features of the CHARGE Association: Support for Involvement of the Neural Crest. *Teratology*, pages 1–6, Nov. 1985.

M. S. Simões-Costa, S. J. McKeown, J. Tan-Cabugao, T. Sauka-Spengler, and M. E. Bronner. Dynamic and Differential Regulation of Stem Cell Factor FoxD3 in the Neural Crest Is Encrypted in the Genome. *PLoS Genetics*, 8(12):e1003142, Dec. 2012.

M. Simões-Costa and M. E. Bronner. Establishing neural crest identity: a gene regulatory recipe. *Development*, 142(2):242–257, 2015. doi: 10.1242/dev.105445.

S. Singh and A. K. Groves. The molecular basis of craniofacial placode development. *Wiley Interdisciplinary Reviews: Developmental Biology*, 5(3):363–376, 2016.

A. Smith, V. Robinson, K. Patel, and D. G. Wilkinson. The epha4 and ephb1 receptor tyrosine kinases and ephrin-b2 ligand regulate targeted migration of branchial neural crest cells. *Current Biology*, 7(8):561 – 570, 1997.

T. N. Snider and Y. Mishina. Cranial neural crest cell contribution to craniofacial formation, pathology, and future directions in tissue engineering. *Birth Defects Research. Part C, Embryo Today : Reviews*, 102(4):324–32, 2014.

T. N. Snider and Y. Mishina. Cranial neural crest cell contribution to craniofacial formation, pathology, and future directions in tissue engineering. *Birth Defects Research. Part C, Embryo today : reviews*, 102(3):324–332, Sept. 2014.

L. Sommer. Generation of melanocytes from neural crest cells. *Pigment Cell & Melanoma Research*, 24(3):411–421, 2011.

E. D. Sperry, E. A. Hurd, M. A. Durham, E. N. Reamer, A. B. Stein, and D. M. Martin. The chromatin remodeling protein chd7, mutated in charge syndrome, is necessary for proper craniofacial and tracheal development. *Developmental Dynamics*, 243(9):1055–1066, 2014.

D. Y. R. Stainier, E. Raz, N. D. Lawson, S. C. Ekker, R. D. Burdine, J. S. Eisen, P. W. Ingham, S. Schulte-Merker, D. Yelon, B. M. Weinstein, M. C. Mullins, S. W. Wilson, L. Ramakrishnan, S. L. Amacher, S. C. F. Neuhauss, A. Meng, N. Mochizuki,

P. Panula, and C. B. Moens. Guidelines for morpholino use in zebrafish. *PLoS Genetics*, 13(10):e1007000, Oct. 2017.

C. D. Stern. The Chick. *Developmental Cell*, 8(1):9–17, Jan. 2005.

B. Steventon, C. Araya, C. Linker, S. Kuriyama, and R. Mayor. Differential requirements of bmp and wnt signalling during gastrulation and neurulation define two steps in neural crest induction. *Development*, 136(5):771–779, 2009.

B. Steventon, R. Mayor, and A. Streit. Neural crest and placode interaction during the development of the cranial sensory system. *Developmental Biology*, 389(1):28 – 38, 2014.

A. Streit. Early development of the cranial sensory nervous system: from a common field to individual placodes. *Developmental Biology*, 276(1):1–15, Dec. 2004.

P. H. Strobl-Mazzulla and M. E. Bronner. A PHD12–Snail2 repressive complex epigenetically mediates neural crest epithelial-to-mesenchymal transition. *The Journal of Cell Biology*, 198(6):999–1010, Sept. 2012.

Strobl-Mazzulla, Pablo Hernan, Sauka-Spengler, Tatjana, and Bronner-Fraser, Marianne. Histone demethylase JmjD2A regulates neural crest specification. *Developmental Cell*, 19(3):460–468, Sept. 2010.

H. L. Szabo-Rogers, L. E. Smithers, W. Yakob, and K. J. Liu. New directions in craniofacial morphogenesis. *Developmental Biology*, 341(1):84–94, May 2010.

I. Takada, M. Mihara, M. Suzawa, F. Ohtake, S. Kobayashi, M. Igarashi, M.-Y. Youn, K.-i. Takeyama, T. Nakamura, Y. Mezaki, S. Takezawa, Y. Yogiashi, H. Kitagawa, G. Yamada, S. Takada, Y. Minami, H. Shibuya, K. Matsumoto, and S. Kato. A histone lysine methyltransferase activated by non-canonical Wnt signalling suppresses PPAR-gamma transactivation. *Nature Cell Biology*, 9(11):1273–1285, Oct. 2007.

L. Taylor, D. F. Carlson, S. Nandi, A. Sherman, S. C. Fahrenkrug, and M. J. McGrew. Efficient TALEN-mediated gene targeting of chicken primordial germ cells. *Development*, 144(5):928–934, Mar. 2017.

M.-A. Teillet, C. Ziller, and N. M. Le Douarin. *Quail-Chick Chimeras*, pages 305–318. Humana Press, Totowa, NJ, 1999. ISBN 978-1-59259-270-8.

P. I. Thakore, A. M. D'Ippolito, L. Song, A. Safi, N. K. Shivakumar, A. M. Kabadi, T. E. Reddy, G. E. Crawford, and C. A. Gersbach. Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nature Methods*, 12(12):1143–1149, Oct. 2015.

E. Theveneau, B. Steventon, E. Scarpa, S. Garcia, X. Trepat, A. Streit, and R. Mayor. Chase-and-run between adjacent cell populations promotes directional collective migration. *Nature Cell Biology*, 15(7):763–772, July 2013.

A. J. Thomas and C. A. Erickson. FOXD3 regulates the lineage switch between neural crest-derived glial cells and pigment cells by repressing MITF through a non-canonical mechanism. *Development*, 136(11):1849–1858, June 2009.

J. Timmer, J. Johnson, and L. Niswander. The use of in ovo electroporation for the rapid analysis of neural-specific murine enhancers. *Genesis*, 29(3):123–132, 2001.

P. A. Trainor. Craniofacial birth defects: The role of neural crest cells in the etiology and pathogenesis of treacher collins syndrome and the potential for prevention. *American Journal of Medical Genetics. Part A*, pages 2984–94, 2010.

G. Trichas, J. Begbie, and S. Srinivas. Use of the viral 2A peptide for bicistronic expression in transgenic mice. *BMC Biology*, 6(1):40–13, 2008.

L. A. Trinh, V. Chong-Morrison, D. Gavriouchkina, T. Hochgreb-Hägele, U. Senanayake, S. E. Fraser, and T. Sauka-Spengler. Biotagging of Specific Cell Populations in Zebrafish Reveals Gene Regulatory Logic Encoded in the Nuclear Transcriptome. *Cell Reports*, 19(2):425–440, Apr. 2017.

M. Uchikawa. Enhancer analysis by chicken embryo electroporation with aid of genome comparison. *Development, Growth & Differentiation*, 50(6):467–474, 2008.

M. Uchikawa, Y. Ishida, T. Takemoto, Y. Kamachi, and H. Kondoh. Functional analysis of chicken sox2 enhancers highlights an array of diverse regulatory elements that are conserved in mammals. *Developmental Cell*, 4(4):509 – 519, 2003.

M. Uchikawa, T. Takemoto, Y. Kamachi, and H. Kondoh. Efficient identification of regulatory sequences in the chicken genome by a powerful combination of embryo electroporation and genome comparison. *Mechanisms of Development*, 121(9): 1145–1158, Sept. 2004.

M. Uchikawa, N. Nishimura, M. Iwafuchi-Doi, and H. Kondoh. *Enhancer Analyses Using Chicken Embryo Electroporation*, pages 191–202. Springer New York, New York, NY, 2017. ISBN 978-1-4939-7216-6.

B. C. Valdez, D. Henning, R. B. So, J. Dixon, and M. J. Dixon. The Treacher Collins syndrome (TCOF1) gene product is involved in ribosomal DNA gene transcription by interacting with upstream binding factor. *PNAS*, 101(29):10709–10714, July 2004.

D. L. C. van den Berg, T. Snoek, N. P. Mullin, A. Yates, K. Bezstarosti, J. Demmers, I. Chambers, and R. A. Poot. An Oct4-centered protein interaction network in embryonic stem cells. *Cell Stem Cell*, 6(4):369–381, Apr. 2010.

J. E. Varley, R. G. Wehby, D. C. Rueger, and G. D. Maxwell. Number of adrenergic and islet-1 immunoreactive cells is increased in avian trunk neural crest cultures in the presence of human recombinant osteogenic protein-1. *Developmental Dynamics*, 203(4):434–447, Aug. 1995.

A. Verloes. Updated diagnostic criteria for CHARGE syndrome: a proposal. *American Journal of Medical Genetics Part A*, 133A(3):306–308, Mar. 2005.

N. Veron, Z. Qu, P. A. S. Kipen, C. E. Hirst, and C. Marcelle. CRISPR mediated somatic cell genome engineering in the chicken. *Developmental Biology*, 407(1):1–7, Aug. 2015.

Visel, Axel, Blow, Matthew J, Li, Zirong, Zhang, Tao, Akiyama, Jennifer A, Holt, Amy, Plajzer-Frick, Ingrid, Shoukry, Malak, Wright, Crystal, Chen, Feng, Afzal, Veena, Ren, Bing, Rubin, Edward M, and Pennacchio, Len A. ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature*, 457(7231):854–858, Feb. 2009.

L. E. L. M. Vissers, C. M. A. van Ravenswaaij, R. Admiraal, J. A. Hurst, B. B. A. de Vries, I. M. Janssen, W. A. van der Vliet, E. H. L. P. G. Huys, P. J. de Jong, B. C. J. Hamel, E. F. P. M. Schoenmakers, H. G. Brunner, J. A. Veltman, and A. G. van Kessel. Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nature Genetics*, 36(9):955–957, Sept. 2004.

D. Vlieghe, A. Sandelin, P. J. De Bleser, K. Vleminckx, W. W. Wasserman, F. van Roy, and B. Lenhard. A new generation of JASPAR, the open-access repository for transcription factor binding site profiles. *Nucleic Acids Research*, 34(Database issue):D95–7, Jan. 2006.

P. Vuorela, S. Ala-Mello, C. Saloranta, M. Penttinen, M. Pöyhönen, K. Huoponen, W. Borozdin, B. Bausch, E. M. Botzenhart, C. Wilhelm, H. Kääriäinen, and J. Kohlhase. Molecular analysis of the CHD7 gene in CHARGE syndrome: identification of 22 novel mutations and evidence for a low contribution of large CHD7 deletions. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, 9(10):690–694, Oct. 2007.

E. A. H. Wanda S Layman and D. M. Martin. Reproductive dysfunction and decreased gnrh neurogenesis in a mouse model of charge syndrome. *Human Molecular Genetics*, 20(16):3138–3150, 2011.

G. G. Wang, C. D. Allis, and P. Chi. Chromatin remodeling and cancer, part II: ATP-dependent chromatin remodeling. *Trends in Molecular Medicine*, 13(9):373–380, Sept. 2007.

M. K. Weng, B. Zimmer, D. Pöltl, M. P. Broeg, V. Ivanova, J. A. Gaspar, A. Sachinidis, U. Wüllner, T. Waldmann, and M. Leist. Extensive transcriptional regulation of chromatin modifiers during human neurodevelopment. *PLoS ONE*, 7(5):e36708, 2012.

J. A. Weston. Sequential segregation and fate of developmentally restricted intermediate cell populations in the neural crest lineage. *Current Topics in Developmental Biology*, 25:133–153, 1991.

S. Whalen, R. M. M. Truty, and K. S. Pollard. Enhancer- promoter interactions are encoded by complex genomic signatures on looping chromatin. *Nature Genetics*, 48(5):1–10, Apr. 2016.

D. Whittaker, K. Riegman, S. Kasah, C. Mohan, T. Yu, B. Sala, H. Hebaishi, A. Caruso, A. Marques, C. Michetti, M. Smachetti, A. Shah, M. Sabbioni, O. Kulhanci, W.-W. Tee, D. Reinberg, M. Scattoni, H. Volk, I. McGonnell, F. Wardle, C. Fernandes, and M. Basson. The chromatin remodeling factor chd7 controls cerebellar development by regulating reelin expression. *The Journal of Clinical Investigation*, 127:874–887, 3 2017.

C. B. Wiese, K. K. Deal, S. J. Ireland, V. A. Cantrell, and E. M. Southard-Smith. Migration pathways of sacral neural crest during development of lower urogenital tract innervation. *Developmental Biology*, 429(1):356–369, Sept. 2017.

M. S. Williams. Speculations on the pathogenesis of CHARGE syndrome. *American Journal of Medical Genetics Part A*, 133A(3):318–325, 2005.

R. M. Williams, U. Senanayake, M. Artibani, G. Taylor, D. Wells, A. A. Ahmed, and T. Sauka-Spengler. Genome and epigenome engineering CRISPR toolkit for *in*

vivo modulation of cis-regulatory interactions and gene expression in the chicken embryo. *Development*, 145(4):dev160333, Feb. 2018.

K. A. Williamson, A. M. Hever, J. Rainger, R. C. Rogers, A. Magee, Z. Fiedler, W. T. Keng, F. H. Sharkey, N. McGill, C. J. Hill, A. Schneider, M. Messina, P. D. Turnpenny, J. A. Fantes, V. van Heyningen, and D. R. Fitzpatrick. Mutations in SOX2 cause anophthalmia-esophageal-genital (AEG) syndrome. *Human Molecular Genetics*, 15(9):1413–1422, May 2006.

T. Woodage, M. A. Basrai, A. D. Baxevanis, P. Hieter, and F. S. Collins. Characterization of the CHD family of proteins. *PNAS*, 94(21):11472–11477, Oct. 1997.

H. Xi, H. P. Shulha, J. M. Lin, T. R. Vales, Y. Fu, D. M. Bodine, R. D. G. McKay, J. G. Chenoweth, P. J. Tesar, T. S. Furey, B. Ren, Z. Weng, and G. E. Crawford. Identification and Characterization of Cell Type–Specific and Ubiquitous Chromatin Regulatory Structures in the Human Genome. *PLoS Genetics*, 3(8):e136–12, 2007.

G. L. Xu, T. H. Bestor, D. Bourc'his, C. L. Hsieh, N. Tommerup, M. Bugge, M. Hulten, X. Qu, J. J. Russo, and E. Viegas-Pequignot. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature*, 402(6758):187–191, Nov. 1999.

T. Yamamori, K. Fukada, R. Aebersold, S. Korsching, M. J. Fann, and P. H. Patterson. The cholinergic neuronal differentiation factor from heart cells is identical to leukemia inhibitory factor. *Science*, 246(4936):1412–1416, Dec. 1989.

J.-S. Yong, D. P. Intriago-Baldeón, and E. W.-F. Lam. FOXD3 controls pluripotency through modulating enhancer activity. *Stem Cell Investigation*, 3(5):17–17, 2016.

T. Yu, L. C. Meiners, K. Danielsen, M. T. Wong, T. Bowler, D. Reinberg, P. J. Scambler, C. M. van Ravenswaaij-Arts, and M. A. Basson. Deregulated fgf and homeotic gene expression underlies cerebellar vermis hypoplasia in charge syndrome. *eLife*, 2:e01305, dec 2013.

J. Zalatan, M. Lee, R. Almeida, L. Gilbert, E. Whitehead, M. La Russa, J. Tsai, J. Weissman, J. Dueber, L. Qi, and W. Lim. Engineering complex synthetic transcriptional programs with crispr rna scaffolds. *Cell*, 160(1):339 – 350, 2015.

K. S. Zaret and J. S. Carroll. Pioneer transcription factors: establishing competence for gene expression. *Genes and Development*, 25(21):2227–2241, Nov. 2011.

G. E. Zentner, W. S. Layman, D. M. Martin, and P. C. Scacheri. Molecular and phenotypic aspects of CHD7 mutation in CHARGE syndrome. *American Journal of Medical Genetics Part A*, 152A(3):674–686, Mar. 2010.

D. Zhang, S. Ighaniyan, L. Stathopoulos, B. Rollo, K. Landman, J. Hutson, and D. Newgreen. The neural crest: a versatile organ system. *Birth Defects Research. Part C, Embryo today : reviews*, 102(3):275–298, Sept. 2014.