UNIVERSITÉ DU QUÉBEC À MONTRÉAL

NEW TOOLS FOR STUDYING *CDX* NEURAL FUNCTIONS

THESIS

PRESENTED

AS PART

OF A MASTER'S IN BIOCHEMISTRY

BY

DANIEL AGUIRRE MARTINEZ

FEBRUARY 2016

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

NOUVEAUX OUTILS POUR L'ETUDE DES FONCTIONS NEURALES DES GENES *CDX*

MÉMOIRE

PRÉSENTÉ

COMME EXIGENCE PARTIELLE

DE LA MAÎTRISE EN BIOCHIMIE

PAR

DANIEL AGUIRRE MARTINEZ

FÉVRIER 2016

ACKNOWLEDGEMENTS

I would like to address my sincerest thanks to the following people who have helped me throughout the last two years of my master's program:

Nicolas Pilon, my supervisor and one of my professors at the university, for giving me the opportunity to learn in his lab. Despite all the piles of books and articles keeping him busy in his office he is always available to talk to.

Ouliana Souchkova for imparting on me valuable technical and scientific knowledge. Especially at the beginning when she had to do all that plus deal with my broken ankle and limited mobility.

All the members of the lab who have helped me in some way or another

My family and friends for their continuing love and support

TABLE OF CONTENTS

ACKNOWLEDGEMENTS i

FIGURE LIST vii

TABLE LIST ix

ABBREVIATION, SIGN AND ACRONYM LIST xi

RÉSUMÉ xiii

ABSTRACT xv

INTRODUCTION 1

CHAPTER I 3
BACKGROUND 3
1.1 Vertebrate embryogenesis 3
1.2 Molecular mechanisms 8
1.2.1 Signals of the posterior neuroectoderm 10
1.3 Gene regulation and NC development 18
1.3.1 Enhancers 18
1.3.2 *Cdx2*NSE 22
1.4 Diseases of NCC origin and neural tube defects 24
1.4.1 Neurocristopathies 24
1.4.2 Neural tube defects 26
1.4.3 Tools for studying neurocristopathies and NTDs 27

CHAPTER II 29
HYPOTHESES AND OBJECTIVES 29
2.1 Hypotheses 29

2.2 Objectives 29

CHAPTER III 31

MATERIALS AND METHODS 31

3.1 Construction of pROSA26-Cdx1, ES targeting and verification 31

3.1.1 Vector construction 31

3.1.2 Targeting ES cells 31

3.1.3 Southern Blotting 32

3.2 Testing the neural specific regulation of *Cdx2*NSE 33

3.2.1 Plasmids 33

3.2.2 Luciferase assay conditions 33

3.2.3 Luciferase assay test and statistical analysis 35

CHAPTER IV 37

TOWARD THE GENERATION OF MICE CARRYING THE CRE-INDUCIBLE *CDX1* TRANSGENE 37

4.1 Introduction 37

4.2 Results 37

4.2.1 Construction of the targeting vector 37

4.2.2 Targeting the *ROSA26* locus 38

CHAPTER V 43

TESTING THE NEURAL SPECIFIC REGULATION OF *CDX2*NSE 43

5.1 Introduction 43

5.2 Results 44

5.2.1 Potential binding activators 44

5.2.2 NSE and potential TFBS sequence homology 46

5.2.3 Directional preference and synergy 49

5.2.4 Testing the reverse NSE 51

CHAPTER VI 53

DISCUSSION 53

6.1 In pursuit of a conditional *Cdx1* overexpression mouse model 53

6.2 Insights into *Cdx2*NSE neural regulation and potential application 56

APPENDIX A 63
EVOLUTIONARY CONSERVATION OF *CDX2*NSE POTENTIAL TFBS 63
APPENDIX B 65
REGULATION OF *PAX3* NCE3 BY CDX2 PROTEINS 65
APPENDIX C 67
*CDX2*NSE AND FGF SIGNALLING 67
REFERENCES 69

FIGURE LIST

Figure Page

1.1 Steps of neurulation....7

1.2 NCC migration throughout four different levels....7

1.3 Neural crest gene regulatory network. 9

1.4 AP axis patterning. 11

1.5 The canonical Wnt signaling pathway. 11

1.6 The FGF signalling pathway. 13

1.7 The spatial and temporal differences of murine *Cdx* gene expression....17

1.8 Types of cis-regulatory elements. 19

1.9 Techniques for identifying and testing enhancers....21

1.10 Identification and testing the Cdx2 neural specific enhancer (NSE). 23

1.11 The enhancer is not entirely neural specific....23

4.1 SacII linearization of pROSA26-Cdx1 39

4.2 Depiction of gene targeting at the *ROSA26* locus. 40

4.3 PCR screen of targeted ES clones. 41

5.1 Potential Lef1 1/Tcf and STEM family binding sites. 45

5.2 Vertebrate homology of the neural specific enhancer 48

5.3 *Cdx2*NSE synergy and directional preference. 50

5.4 Fragmentation results of *Cdx2*NSEreverse. 52

TABLE LIST

Table Page

1.1 Diversity of Neural Crest Cell derivatives. 8

3.1 Primers for PCR screen. 32

3.2 P-value guide 35

3.3 Oligos to construct *Cdx2*NSE reverse strand fragments. 36

ABBREVIATION, SIGN AND ACRONYM LIST

AP	Anterior-Posterior
BMP	Bone morphogenic protein
Cdx	Caudal-related homeobox
CNS	Central nervous system
DLHP	Dorsal-lateral hinge points
DV	Dorsal-Ventral
EMT	Epithelial-mesenchymal transition
EGFP	Enhanced green fluorescence protein
FGF	Fibroblast growth factor
GRN	Gene regulatory network
IRES	Internal ribosome entry site
LEF1	Lymphoid enhancer-binding factor 1
MEF	Mouse Embryo Fibroblast
MHP	Median hinge points
NC	Neural crest

NCC	Neural crest cell
NCE	Neural crest enhancer
NP	Neural Plate
NPB	Neural Plate Border
NSE	Neural Specific Enhancer
NT	Neural Tube
NTD	Neural Tube Defects
PNS	Peripheral nervous system
SOX2	SRY (sex determining region Y)-box 2
TCF	T-cell specific transcription factor
TF	Transcription factor
TFBS	Transcription factor binding site
Wnt	Wingless-related integration site

RÉSUMÉ

Le système nerveux est un composant fondamental chez les vertébrés qui implique un réseau complexe de gènes régulateurs (GRN). Sa mise en place commence avec l'induction du neuroectoderme par l'activité des signaux Wnt, FGF et BMP le long de l'axe antéro-postérieur nouvellement formé. Plusieurs signaux transcriptionnels et morphogéniques coopèrent pour induire l'invagination du neuroectoderme ventralement afin de former le tube neural et la crête neurale qui donneront naissance, respectivement, au système nerveux central et au système nerveux périphérique. Le neuroectoderme postérieur possède son unique lot de molécules de signalisation et toute perturbation dans les évènements précoce de son développement entraine un groupe de maladies et malformations congénitales distinct. L'utilisation de modèles transgéniques murins est une façon idéale de comprendre les activités transcriptionnelles et morphogéniques dans cette région. Nous proposons deux modèles basés sur les facteurs de transcription de la famille *Cdx* – essentiels pour le développement postérieur de l'embryon, et des acteurs majeurs dans le développement de la crête neurale (Sanchez-Ferras et al., 2014; Sanchez-Ferras et al., 2012). Le premier serait un modèle pour l'anomalie qui pourrait être entrainée par la surexpression conditionnelle du gène *Cdx1* et nous permettrait ainsi de mieux comprendre le rôle de ce gène dans les cancers dérivés des cellules de la crête neurale. A ce jour, un criblage par PCR nous a permis d'identifier des cellules souches positives pour l'intégration d'une cassette *Cdx*1 mais une confirmation par southern blot est nécessaire. Concernant la deuxième souris transgénique, elle représente un modèle unique pour étudier les évènements précoces dans le neuroectoderme postérieur. En effet, elle est basée sur l'utilisation de l'enhancer (NSE) de *Cdx2* qui permet de générer une lignée transgénique avec une induction spécifique au tube neural via la technique Cre-LoxP. Cependant, *Cdx2*NSE s'est révélée n'être que partiellement active dans le mésoderme. Cette activité mésodermale étant mineure, il serait possible de retirer les éléments non-neuraux de NSE afin d'obtenir une activité entièrement neurale. Pour déterminer la nature de l'activité des différents éléments de NSE, des essais luciférase ont été effectués par co-transfection du facteur pro-neural Sox2 et d'une protéine de fusion Lef1-βcaténine agissant comme effecteur de la voie Wnt, ces deux derniers étant importants pour le maintien de l'identité neuroectodermale postérieure. Les résultats ont montré que NSE est régulé par Wnt et Sox2 de manière synergistique, mais les éléments non-neuraux restent à identifier.

Mots-clés : GRN, neuroectoderme, postérieur, développement, Cdx, régulation, outils

ABSTRACT

The nervous system is a fundamental feature of vertebrates whose formation involves a complex gene regulatory network (GRN). The process can be said to commence with the induction of the neuroectoderm by Wnt, FGF and BMP activity along the newly formed anterior-posterior axis. Several transcriptional and morphogenic signals in concerted action internalize the neuroectoderm ventrally where it becomes the neural tube and neural crest, the precursors of the central and peripheral nervous system, respectively. The posterior end of the neuroectoderm has its own unique set of signalling molecules and disruptions in the early events in this region engender a distinct set of congenital diseases and malformations. One way to understand the transcriptional and morphogenic activity in this posterior region is through the use of mouse models. Described are two proposed mouse systems based on the genes of the Cdx family of transcription factors—important for posterior embryonic development, and key players in neural crest development (Sanchez-Ferras et al., 2014; Sanchez-Ferras et al., 2012). The first would serve as a model for the potential pathogenesis brought on by *Cdx1* conditional overexpression, in particular its potential role in the oncogenesis of neural crest cell derived cancers. So far, positive embryonic stem cell clones with the *Cdx1* overexpression cassette have been identified via a preliminary PCR screen but require confirmation by southern blot. The second mouse model would take advantage of a modified neural specific enhancer of *Cdx2* to generate a truly neural specific Cre driver line for the posterior neuroectoderm, a unique and useful model for studying the early events in the posterior neuroectoderm. Initially considered as entirely neural specific (Wang and Shashikant, 2007), a Cre mouse line driven by the *Cdx2*NSE later revealed partial activity in the mesoderm as well (Coutaud and Pilon, 2013). Since gene expression was minor in the mesoderm, it seems conceivable that the mesoderm expression can be removed, resulting in true neural specific expression. To do this the non-neural specific elements in *Cdx2*NSE sequence have to be identified and either removed or made non-functional. To this end, luciferase assays were carried out via co-transfecting neural Sox2 and Wnt effectors Lef1-βcatenin. Both are important for maintaining posterior neuroectoderm identity. Results show the NSE is regulated by the synergy of Wnt activity and Sox2. However, the non-neural elements have yet to be identified.

Keywords: GRN, neuroectoderm, posterior, development, Cdx, regulation, tools

INTRODUCTION

A hallmark of vertebrate embryogenesis is the development of the nervous system. The process begins with the induction of the neuroectoderm across the midline. The new structure is not only physically and behaviourally unique from the surrounding epidermal ectoderm but has itself two distinct domains: the neural plate border (NPB), and the neural plate (NP). The NPBs, located on the lateral edges of the neuroectoderm, contain the precursors of the neural crest cells (NCC). The neural plate, located between the NPBs, encompasses a relatively large piece of neuroepithelium that becomes the nascent central nervous system via the formation of the neural tube (NT). The neuroectoderm eventually invaginates ventrally to form the NT and NCC population. The NCCs are a multipotent migratory group of cells that can become peripheral neurons, glia, pigment cells, cartilage cells and others (Garnett et al., 2012; Perris, 1997; Stuhlmiller and Garcia-Castro, 2012).

The transition from ectoderm to neural plate to multipotent neural crest cells requires a finely tuned neural crest-gene regulatory network (NC-GRN) whose interactions are constantly being updated, as they are made known (Amore et al., 2003; Hinman et al., 2003; Oliveri et al., 2003; Van Otterloo et al., 2013). Work in revising and updating this regulatory gene network is valuable not just to improve our understanding of neurogenesis in general but how changes to this network can affect development. One way to shed some light on this gene network is through transgenic mouse models. Mouse systems can provide *in vivo* evidence on NC-GRN processes both spatially and temporally. Two proposed mouse models to be described in this memoire have their origins in the caudal end of development, as they are based on the genes encoding the caudally-restricted Cdx family of transcription factors. Cdx transcription

factors are known for their importance in posterior embryonic development, but have just recently been implicated in the initial steps of the caudal NC-GRN, acting as mediators, inducers and co-operators of known players of the NC regulatory circuit (Sanchez-Ferras et al., 2014; Sanchez-Ferras et al., 2012).

The first mouse model would elucidate the effects of conditionally overexpressing *Cdx1*. Overexpression targeting NCC development is believed to result in the pathogenesis of NCC derived cancers, such as neuroblastoma. So far, potentially positive embryonic stem cell clones possessing the transgene construct have been identified via a preliminary PCR screen but verification by southern blot has of yet been inconclusive. Once verified, microinjection and rearing of transgenic mice capable of inducible *Cdx1* overexpression could proceed. The second model is envisioned to serve as a Cre driver line for Cre-LoxP systems aiming to study the early events of the posterior neuroectoderm. Developed in the lab by Coutaud and Pilon (2013), *Cdx2*NSE-Cre was designed to provide Cre expression solely in the posterior neuroectoderm. Initially characterized as a neural specific enhancer (NSE) by Wang and Shashikant (2007) it was later revealed to be partially active in non-neural mesodermal tissue (Coutaud and Pilon, 2013). To address this problem, it was imperative that the non-neural characteristics of the enhancer be identified and removed or made non-functional thereby preserving a true NSE capable of being employed as a Cre driver line for true neural specific induction in Cre-LoxP systems. To test the absence or presence of a neural identity, luciferase assays were carried out on a *Cdx2*NSE-Luc reporter, co-transfecting neural Sox2 and Wnt signalling effectors Lef1-βcatenin. Wnt is important for determining posterior structures and regulating *Cdx* genes (Pilon et al., 2006; Prinos et al., 2001), and Sox2 is a neural marker of the neuroectoderm (Graham et al., 2003; Uchikawa et al., 2011). So far, results show the NSE is regulated in synergy by combined Wnt activity and Sox2. However, we have not been able to identify the specific areas of the NSE architecture responsible for the non-neural activity.

CHAPTER I

BACKGROUND

1.1 Vertebrate embryogenesis

Following fertilization and cleavage, the third and fourth major events in vertebrate development are gastrulation and neurulation. Prior to gastrulation, the developing embryo is composed of a spherical layer of cells. Gastrulation transforms the single layer into three distinct germ layers. Known as the ectoderm, mesoderm, and endoderm, each of the three germ layers establishes the framework of all future tissues. Moreover, gastrulation begins to shape the anterior-posterior (AP) axis and dorsal midline (Tam and Behringer, 1997). These axes are defined by the formation of the mesodermal derived notochord (Purves and Williams, 2001).

Neurulation (Figure 1.1a) begins with the induction of the neural plate along the AP axis, followed by specification of the neural plate borders (NPB). Housed in the NPB are the precursors of the neural crest cells (NCC). The NPB separates the non-neural ectoderm (NNE) from the neuroectoderm. Next, the neuroectoderm begins to bend inward creating neural folds at the lateral edges. Bending continues inward as the neural folds elevate; such is the degree of the bending that the resulting U-shape is called the neural groove. As the groove lowers the neural folds rise, meet, fuse and close the loop. The closed loop, now a cylinder of neuroepithelial cells known as the neural tube (NT), is freed from the overlying ectoderm. The NT is completed first in the middle of the AP axis, next extending cranially and caudally; in birds and

mammals the NT forms first in cranial then truncal levels (Figure 1.1b) (Colas and Schoenwolf, 2001; Duband, 2010).

At the same time the NT forms, NCCs are freed from what used to be the NPB and they begin to migrate. NCCs exhibit an epithelial to mesenchymal transition (EMT) allowing them to delaminate from the neuroepithelium, and migrate on pathways that will specify their lineage. Delamination and migration occurs in a rostral to caudal wave, coinciding with NT formation, with the wave of migration ultimately filling up four levels: cranial, cardiac/vagal, truncal and sacral (Figure 1.2) (Huang and Saint-Jeannet, 2004; Mason, 2007). In the trunk, NCCs migrate around the NT and through, and sometimes between, somites* and in the head through rhombomeres† (Ghysen, 2003; Gomez et al., 2008; Guthrie and Lumsden, 1991). Ventral regions are filled-up first before more dorsal regions (Weston and Butler, 1966) and depending on where they end up, NCCs can give rise to a diverse set of derivatives of mesenchymal, neuronal, secretory or pigmented identity (Table 1.1) (Simoes-Costa and Bronner, 2015; Smith and Schoenwolf, 1997).

There is still some controversy as to whether NCCs are fate-restricted prior to migration, during migration or if they maintain multipotency (Jessen and Mirsky, 2005; Krispin et al., 2010; Mayor and Theveneau, 2013; McKinney et al., 2013). *In vitro* experiments have demonstrated the multipotent capacity of NCCs to become multiple derivatives such as neurons, osteoblasts and melanocytes (Dupin et al., 2010; Dupin and Sommer, 2012). Moreover, recent advances in *in vivo* cell-labeling and time-lapse imaging show that pre-migratory NCCs from any location along the dorsal

* Segments of paraxial mesoderm that will form the skeletal muscles, vertebraes and ribs

† Neuromere segments that become the future hindbrain

NT may contribute to multiple NC targets, and that microenvironment cues along the migration pathway may be responsible for specifying NCC fate (McKinney et al., 2013). There is also recent evidence suggesting that most migratory NCCs are multipotent and only a few in the population are fate-restricted (Baggiolini et al., 2015).

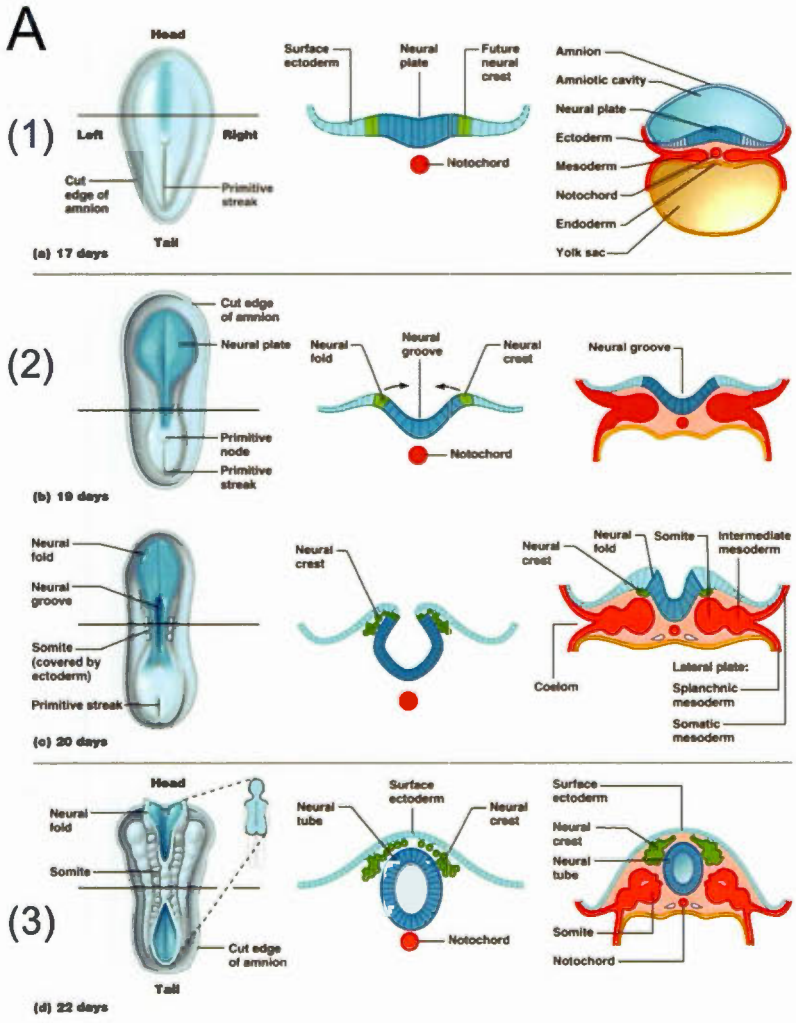

A
(1)
Head
Left
Right
Cut edge of amnion
Primitive streak
Tail
(a) 17 days
Surface ectoderm
Neural plate
Future neural crest
Notochord
Amnion
Amniotic cavity
Neural plate
Ectoderm
Mesoderm
Notochord
Endoderm
Yolk sac
(2)
Cut edge of amnion
Neural plate
Primitive node
Primitive streak
(b) 19 days
Neural fold
Neural groove
Neural crest
Notochord
Neural groove
Neural fold
Neural groove
Somite (covered by ectoderm)
Primitive streak
(c) 20 days
Neural crest
Neural crest
Neural fold
Somite
Intermediate mesoderm
Coelom
Lateral plate:
Splanchnic mesoderm
Somatic mesoderm
(3)
Head
Neural fold
Somite
Cut edge of amnion
Tail
(d) 22 days
Neural tube
Surface ectoderm
Neural crest
Notochord
Surface ectoderm
Neural crest
Neural tube
Somite
Notochord
B
Chick cranial region (Midbrain)
Chick truncal region (Brachial))
24 h (3 s)
28 h (5 s)
32 h (7 s)
36 h (10 s)
42 h (13 s)
64 h (28 s)
42 h (13 s)
48 h (17 s)
52 h (20 s)
54 h (23 s)
68 h (30 s)
80 h (38 s)

Figure 1.1 Steps of neurulation.
(A) Dorsal and transverse view of human neurulation. Neurulation entails at least 3 spatially and temporally distinct stages: (1) formation and shaping of the neural plate (NP); (2) bending of the NP into the neural groove; (3) closure of the neural groove into the neural tube (NT) and subsequent neural crest cell migration. Image retrieved from Marieb and Hoehn (2007).
(B) Chick cranial and truncal temporal differences in NT closure. Birds and mammals exhibit NT closure temporal differences depending on the region along the midline. Yellow: presumptive NCCs; Green: delaminating NCCs; Blue: migrating NCCs. Image retrieved from Duband (2010)

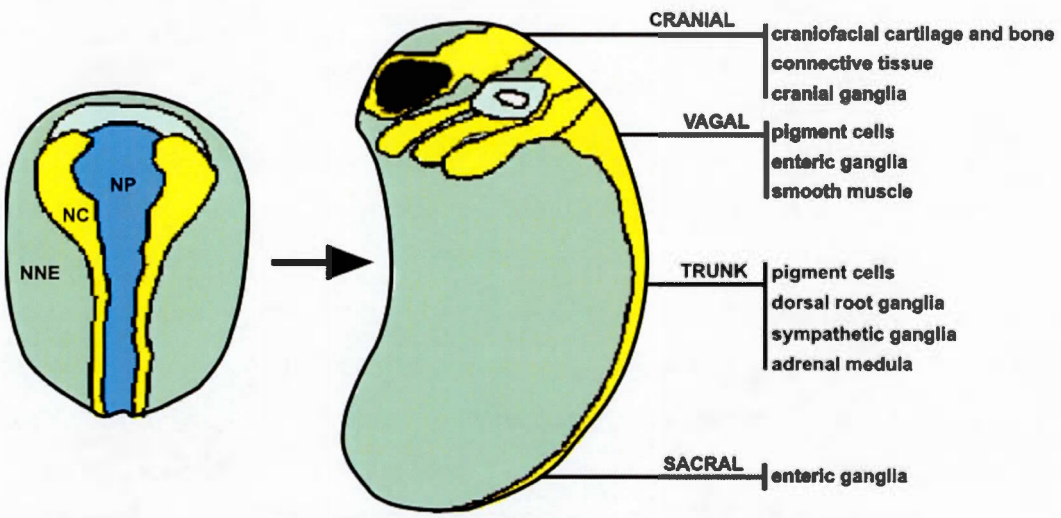


Figure 1.2 NCC migration throughout four different levels.
Migration is the first step in NCCs adopting different lineages. In a process known as epithelial to mesenchymal transition (EMT), presumptive NCCs exit the neuroepithelium and take on different migration pathways each one leading to a change in gene expression and thereby a change in cell type (McKinney et al., 2013). Grey: NNE non-neural ectoderm; Yellow: NC neural crest; Blue: NP neural plate. Image retrieved from Huang and Saint-Jeannet (2004)

Table 1.1 Diversity of Neural Crest Cell derivatives.

Neural crest cell derivatives			
Mesenchymal cells	**Neuronal Cells**	**Secretory Cells**	**Pigmented Cells**
Chondroblasts	Sensory neurons	Chromaffin cells	Melanocytes
Osteoblasts	Cholinergic neurons	Calcitonin-producing cells	
Fibroblasts	Adrenergic neurons	Parafollicular cells	
Odontoblasts	Satellite cells		
Cardiac mesenchyme	Schwann cells		
Myoblasts	Glial cells		
Adipocytes			

Adapted from Simoes-Costa and Bronner (2015).

1.2 Molecular mechanisms

From the transformation of ectoderm to neuroectoderm, to the changing expression profiles of migrating NCCs, there are a multitude of morphogenic signals, cues and gradients, both spatial and temporal that are responsible for these very physical and behavioural changes. This complex network is built on effective communication from one cell to another, relaying several specifying signals at once. This complex network is known as the neural crest gene regulatory network (NC-GRN) (Meulemans and Bronner-Fraser, 2002; Oliveri et al., 2003). Neurulation and NCC migration/specification rely on the NC-GRN (Figure 1.3). The process begins with the induction of the neural plate via the Wnt, FGF and BMP signalling pathways (Huang and Saint-Jeannet, 2004). It is the combination of these three signals and their antagonists that create the gradients responsible for cell specification and morphogenesis (Niehrs, 2004). The specifics as to how these three signalling pathways work together in a concerted manner to induce the NPB is not fully understood and is actively being researched (Garnett et al., 2012; Monsoro-Burq et al., 2005).

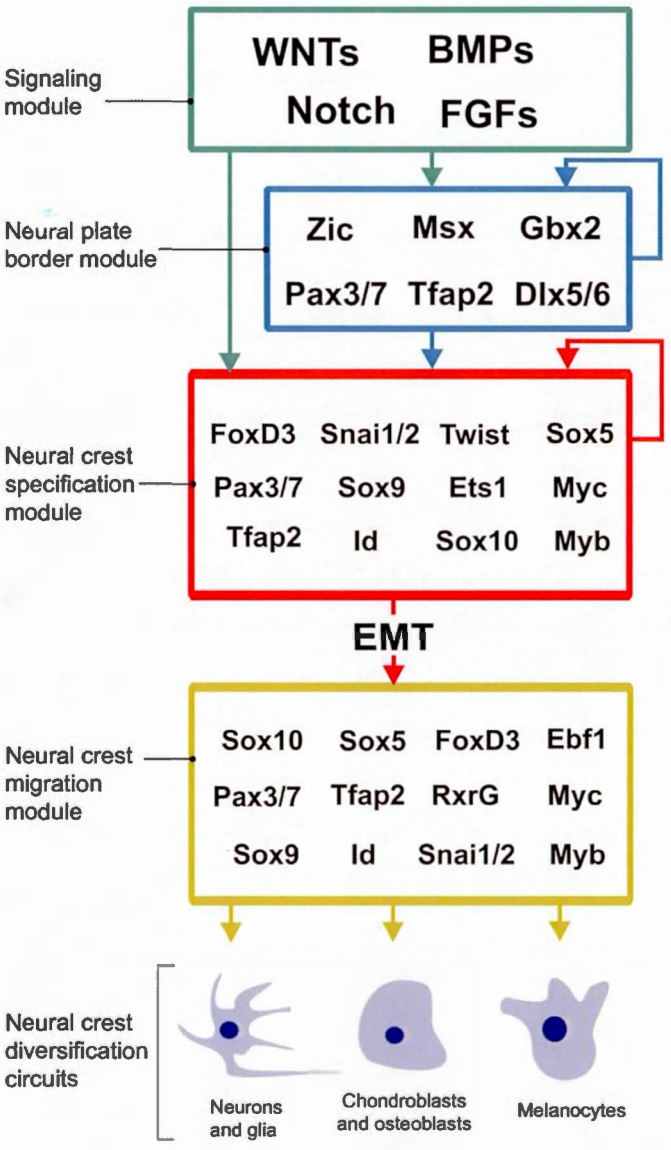


Figure 1.3 Neural crest gene regulatory network.
Through different molecular signals, both temporal and spatial dependent, the NC-GRN is responsible for the process of NC development, from the induction of the NP to the diversification of NC targets. Image retrieved from Simoes-Costa and Bronner (2015).

1.2.1 Signals of the posterior neuroectoderm

The posterior neuroectoderm receives a different concentration, combination and timing of signals than the anterior level. The outcome is a later onset of the NT and NCCs in the posterior region as well as different NCC lineages. In summary, anterior structures are patterned first while posterior ones occur later (Durston, 2015).

The patterning of the anterior-posterior (AP) axis is regulated in part by Wnt and FGF, as well as Nodal and Retinoic acid morphogens and their antagonists (McGrew et al., 1997). Higher Wnt and FGF concentrations specify posterior fates, and their antagonism anterior fates (Figure 1.4) (Kiecker and Niehrs, 2001; Kudoh et al., 2002). Meanwhile, BMP and its antagonists* pattern the dorsal-ventral (DV) axis (Marchant et al., 1998; Patten and Placzek, 2002; Smith and Harland, 1992). Higher BMP concentrations specify ventral fates and lower levels dorsal fates (Endo et al., 2002; Little and Mullins, 2006). Again, all three —Wnt, FGF and BMP— morphogens are needed to induce the formation of the NPB (home of the future NCCs) but Wnt and FGF signals are mostly† posteriorizing; BMP mostly‡ ventralizing (Hendrickx et al., 2009; Tuazon and Mullins, 2015).

---

* such as Noggin (Nog), Follistatin (Flst), and Chordin (Chd)

† Wnt and FGF signaling are crucial for AP patterning but they can be involved in DV patterning as they can affect regulation of BMP activity

‡ BMP is crucial for DV patterning but isoform BMP4 has posteriorizing effects. In all, there is a trend where Wnt and FGF : posteriorizing and BMP : ventralizing, but overlap may occur

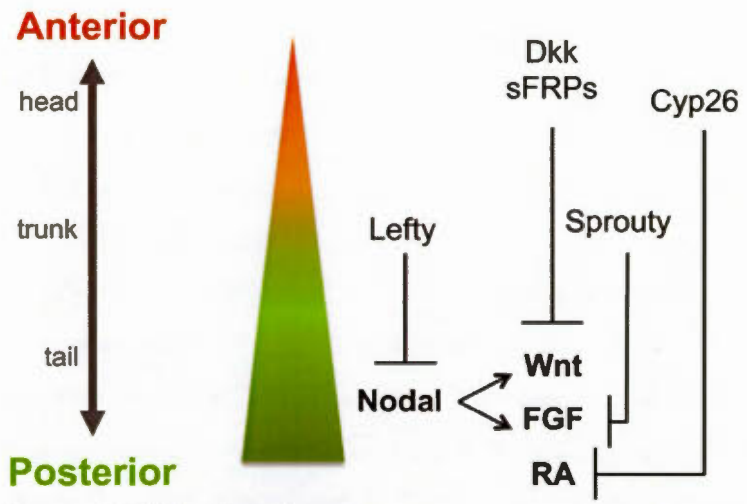


Figure 1.4 AP axis patterning.
Higher concentrations of Wnt FGF Nodal and RA activity specify posterior fates; antagonists to these posteriorizing signals specify anterior fates. Image retrieved from Tuazon and Mullins (2015)

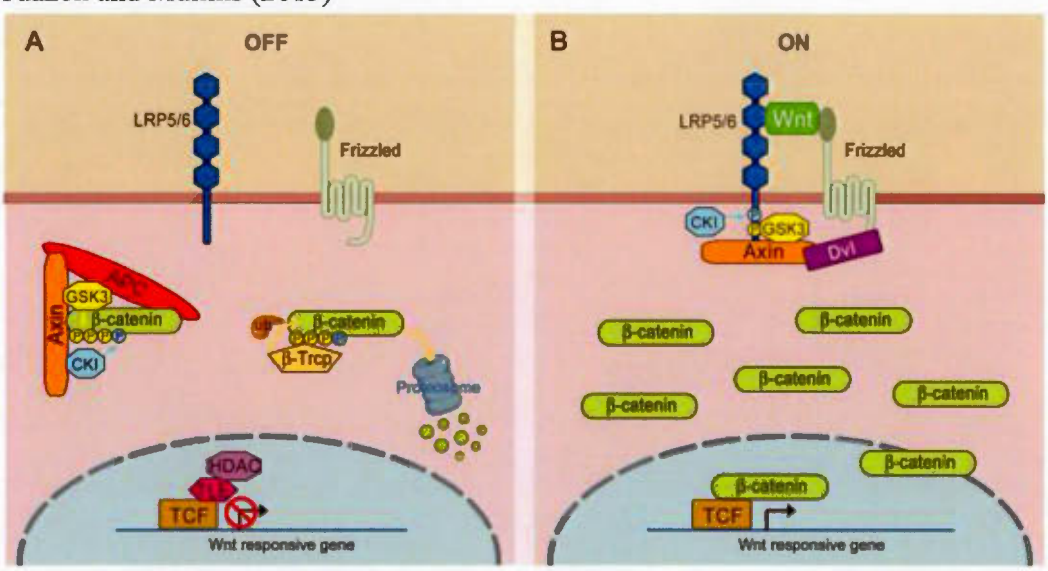


Figure 1.5 The canonical Wnt signalling pathway.
(A) With no Wnt ligand present, the β-catenin destruction complex is active and thus no β-catenin is freely available to mediate Wnt responsive gene transcription. (B) Upon introduction of a Wnt ligand to the Frizzled receptor, the Dishevelled proteins are activated and impede the destruction of β-catenin. This causes cytoplasmic levels of β-catenin to rise to where they are then able to enter the nucleus and associate with Lef1/Tcf family of DNA-binding proteins and activate target gene transcription. Image retrieved from MacDonald et al. (2009)

1.2.1.1 Wnt and FGF signalling

The canonical Wnt signalling pathway begins with the binding of Wnt ligands to Frizzled receptors, which in turn stops the β-catenin destruction complex. An increasing concentration of β-catenin in the cytoplasm favours its translocation into the nucleus where it can then associate with transcription factors of the Lef1/Tcf family and activate target gene expression (Figure 1.5) (Angers and Moon, 2009; MacDonald et al., 2009; Petersen and Reddien, 2009). This modulation of gene expression by the Wnt pathway is implicated in most stages of embryogenesis throughout various phyla including deuterostomes, protostomes, and pre-bilaterians (Petersen and Reddien, 2009; Stuhlmiller and Garcia-Castro, 2012). In vertebrates, the Wnt pathway is responsible for cell proliferation, cell fate, and body AP axis patterning (Hikasa and Sokol, 2013). Particular to AP neural patterning is how Wnt specifies caudal CNS cell fates (Tuazon and Mullins, 2015). Inversely, anterior levels require antagonists of Wnt activity (Houart et al., 2002). For the most part, posterior markers are tied to an increase in Wnt activity whereas a decrease in Wnt activity increases anterior markers (Kudoh et al., 2002; McGrew et al., 1997; Petersen and Reddien, 2009).

Like the Wnt pathway, the FGF pathway is important in establishing the vertebrate AP axis and specifying posterior cell fates (Figure 1.4;1.6) (Tuazon and Mullins, 2015). Suppressing FGF blocks posterior tissue development; overexpressing FGF blocks anterior tissue development (Draper et al., 2003; Isaacs et al., 1994; Stuhlmiller and Garcia-Castro, 2012). As such, both Wnt and FGF signals inhibit anterior gene expression (Kudoh et al., 2002). Concerning their origin, Wnt signals come from the non-neural ectoderm and the paraxial mesoderm; FGF only comes from the paraxial mesoderm (Huang and Saint-Jeannet, 2004). NPB induction may require both Wnt and FGF signals, but it is still unclear if FGF acts directly to induce

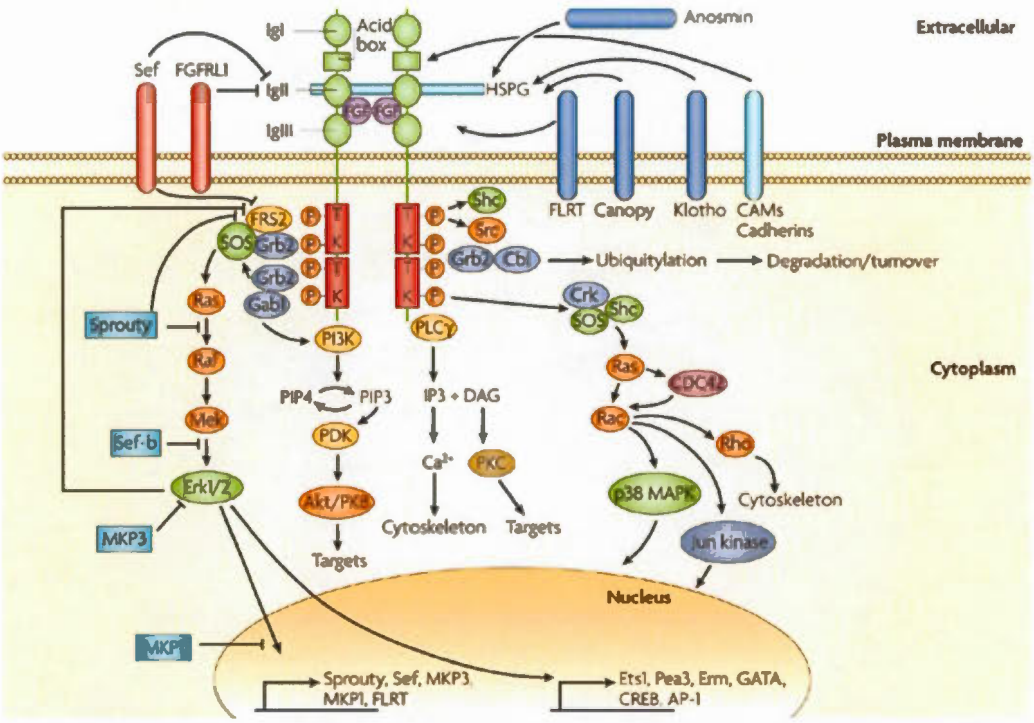


Figure 1.6 The FGF signalling pathway.
FGF binding to FGF receptors leads to the recruitment of Grb2 and Ras among other effectors to activate the Erk1/MAP kinase pathway which in turn phosphorylates a diverse set of transcription factors initiating their regulatory activity implicated in cell growth, migration, and morphogenesis. In AP patterning, antagonists of the FGF pathway, such as Sprouty, work to inhibit posteriorizing fates, thus prompting anterior ones. Image retrieved from Mason (2007).

the NPB or indirectly by activating the Wnt pathway (Garnett et al., 2012). Several studies by Hong et al. (2008) have suggested that Wnt originating from the paraxial mesoderm is activated by FGF, indicating that FGF activates mesodermal Wnt which in turn helps induce NCC formation.

Together, through overlap, direct or indirect methods, Wnt, FGF and BMP signalling induce the NC precursors at the NPB. The first phase involves Wnt and FGF to activate the expression of border specifiers such as *Msx1/2 Pax3/7* and *Zic*. The second phase involves Wnt, BMP as well as BMP-antagonizing Notch signalling to further specify the NCC, by stimulating the expression of *Snail2*, *FoxD3*, and *Sox9/10* (NC-GRN Figure 1.3) (Huang and Saint-Jeannet, 2004; Tuazon and Mullins, 2015).

1.2.1.2 Sox2

Important for the development of the neural primordia as well are the Sox-B1 transcription factors: Sox1, Sox2 and Sox3 (Uchikawa et al., 2011). Sox2 expression in particular is the most extensive, with full expression across the neural primordia (Okuda et al., 2010). In fact, Sox2 is the most definitive marker of the early neural plate (Papanayotou et al., 2008; Pevny and Nicolis, 2010; Rex et al., 1997). From embryo to adult, Sox2 plays a role in maintaining neural progenitor populations (Brazel et al., 2005; Ellis et al., 2004). Constitutively expressing Sox2 suppresses neuronal differentiation; inhibiting Sox2 results in early neuronal differentiation (Graham et al., 2003). Suppressing neuronal differentiation in the early neuroectoderm is essential for maintaining the neural plate identity (Kishi et al., 2000). In the posterior levels, during neural plate development, Wnt and FGF signals work together to activate *Sox2* expression (Takemoto et al., 2006).

1.2.1.3 *Cdx* genes

Caudal *Cdx* genes are also implicated in the development of the posterior NP and in AP patterning (Marom et al., 1997), as well as mediating the closure of the NT (Savory et al., 2011a) and players of the truncal NC-GRN (Sanchez-Ferras et al., 2014; Sanchez-Ferras et al., 2012). In addition, they play a significant role in gut development (Beck and Stringer, 2010; Silberg et al., 2000).

Cdx genes are related to the caudal *Drosophila* gene (*cad*) (Barad et al., 1988; Mlodzik et al., 1985). There are three *Cdx* genes: *Cdx1*, *Cdx2*, and *Cdx4*. During embryogenesis, the three are expressed in the caudal regions, occupying all tissues surrounding the primitive streak (Houle et al., 2003a). Their timing and location around this area varies; the *Cdx1* expression domain has the most rostral reach, followed by *Cdx2* (Figure 1.7). *Cdx2* is the earliest expressed, E3.5 at the trophoectoderm. *Cdx2*-null mice die at this stage because they fail to implant (Strumpf et al., 2005). In embryonic tissues, *Cdx1/4* are expressed around E7.5 before *Cdx2* (E8.5). Expression of all three attenuates as time goes on and recedes caudally until *Cdx4* expression is gone at E10.5 and *Cdx1/2* expression remain localized in the gut where they continue throughout life (Lohnes, 2003; Silberg et al., 2000). In AP patterning, Cdx1/2 are functionally similar despite sequence variability (Savory et al., 2009b; van den Akker et al., 2002). Knockout studies show that *Cdx1* is important for AP patterning (Subramanian et al., 1995); *Cdx2* is important for AP patterning and gut development (Beck and Stringer, 2010; Chawengsaksophak et al., 2004) and *Cdx4* has subtle importance in AP patterning (van Nes et al., 2006).

The three Cdx proteins regulate *Hox* gene expression, both directly (Beck et al., 1995; Marom et al., 1997; Subramanian et al., 1995) and indirectly (Savory et al., 2009a). In vertebrates, *Hox* genes control axial regionalization, as well as the subdivision of the nascent vertebrae (Iimura et al., 2009). Studies suggest that *Cdx* genes regulate *Hox* genes by conveying the posteriorizing Wnt, FGF, and RA signals (Bel-Vialar et

al., 2002; Houle et al., 2003b; Keenan et al., 2006; Lohnes, 2003; Pilon et al., 2006; Pilon et al., 2007; Shimizu et al., 2005). Wnt in particular seems to be important in the process, with combinations of Wnt and FGF (Keenan et al., 2006) and Wnt and RA (Pilon et al., 2007) demonstrating the significant contribution of Wnt signals in inducing *Cdx* expression. Moreover, *Cdx* genes have been found to mediate Wnt signalling in specifying posterior morphogenesis in vertebrates (Sanchez-Ferras et al., 2012; Shimizu et al., 2005; Zhao et al., 2014). Cdx1 can regulate its own promoter, and significantly more so with Lef1 (a Wnt effector), suggesting the existence of Wnt-Cdx joint regulatory complex (Beland et al., 2004). *Cdx4* can also be regulated by Wnt signalling (Pilon et al., 2006) as well as by Cdx2 (Savory et al., 2011b).

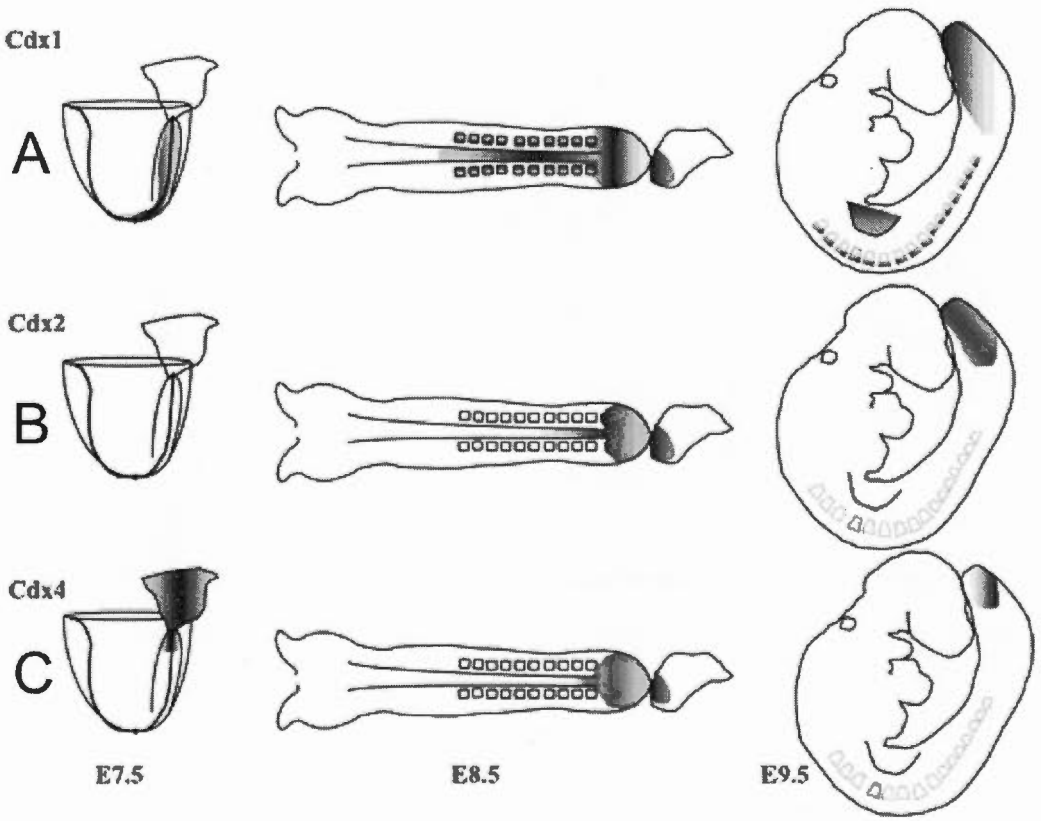


Figure 1.7 The spatial and temporal differences of murine *Cdx* gene expression. **(A)** *Cdx1* expression begins at E7.5 at the ectoderm and mesoderm of the primitive streak. At E8.5, expression extends to the neural plate, paraxial mesoderm, hindgut and tailbud; At E9.5, expression regresses caudally, and lingers in somites and presumptive dermamyotome; at E12.5,, expression is limited to the gut where it persists (Houle et al., 2003a; Meyer and Gruss, 1993; Silberg et al., 2000). **(B)** *Cdx2* embryonic expression begins at E8.5 in the neural plate, neural tube, part of the notochord, the hindgut and all tissues of the tailbud; at E9.5, expression regresses caudally in the NT, NP and notochord, hindgut and all tissues of the tailbud more toward the tail bud; expression is limited to the gut at E12.5 where it persists (Beck et al., 1995; Silberg et al., 2000). **(C)** *Cdx4* expression begins at E7.5 near the posterior end of the primitive streak; at E8.5-9.5, expression regresses caudally in the mesoderm and hindgut endoderm; at E10.5, expression ends (Gamer and Wright, 1993; Lohnes, 2003). Image retrieved from Houle et al. (2003a)

1.3 Gene regulation and NC development

NC development relies on a feed-forward system of regulatory events known as the gene regulatory network (Figure 1.3) (Meulemans and Bronner-Fraser, 2004). Vital to any GRN is proper gene regulation by cis-regulatory regions. These regions are so important that it has been suggested that mutations in pre-vertebrate cis-regulatory regions were critical for NC evolution, and by extension the evolution of the NC-GRN (Jandzik et al., 2015; Van Otterloo et al., 2013).

1.3.1 Enhancers

Enhancers are small cis-regulatory elements of around 200-500bp that can be up to 1.5 Mb downstream, upstream or even intronic of their target gene; (Figure 1.8) (Epstein, 2009; Rada-Iglesias et al., 2013). Their function is to enhance target gene expression by activating the promoter of the target gene (Pennacchio et al., 2013). Transcription factors bind to specific regulatory motifs on the enhancer and mediate activation of the target promoter. In nature, an activated enhancer elicits a conformation of the chromatin structure so that the enhancer can loop near the target promoter and mediate its activation (Kranz et al., 2011).

During development enhancers play a role in cellular growth, differentiation and migration by relaying the activator signals from growth factors and transcription factors (Howard and Davidson, 2004; Kranz et al., 2011). In the context of the NC-GRN, for example, enhancers of NPB specifiers, such as Msx, Pax3/7, Zic1, Dlx3/5 would be at the receiving end of effectors of Wnt, FGF, BMP and Notch signalling. Then, these NPB specifier transcription factors, along with other signals, would be at the delivery end of the enhancers of NC specifiers, such as Zic3, Sox9, Foxd3, Sox10, Snai2, and Twist, and so on until completion of NCC differentiation.

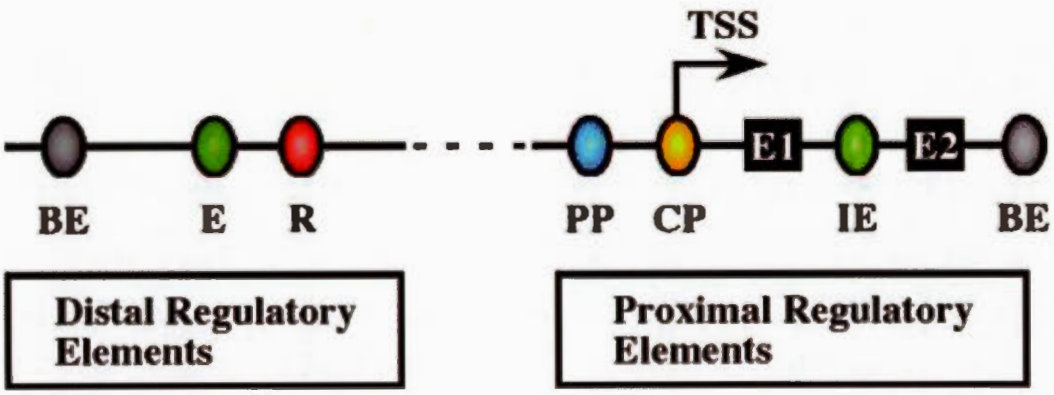


Figure 1.8 Types of cis-regulatory elements.
From left to right, BE: boundary element; E: enhancer; R: repressor (the counterpart of the enhancer—instead of promoting the activation of a target promoter they suppress activation) PP: proximal promoter; CP: core promoter; TSS: transcription start site; IE: intronic enhancer. E1/2 represent exons. Image retrieved from Epstein (2009)

Also of note regarding the current NC-GRN, as it relates to enhancer activity, is the absence of known overlying epigenetic regulation needed for enhancer activity to occur in the first place. In fact, several epigenetic modifiers regulating NC development have been identified, and constitute a fairly recent area of active research (Hu et al., 2014; Liu and Xiao, 2011).

1.3.1.1 Approaches for identifying and testing enhancers

According to Simoes-Costa and Bronner (2013), there are three main methods used in identifying enhancers (Figure 1.9). All three require: the construction of the putative enhancer sequence with a minimal promoter to drive the expression of a reporter, and the testing of the putative enhancer via activation through cell transfections, or *in vivo* stable or transient inducible animal models.

The first method requires inserting several non-coding sequences, one by one, into a reporter construct in the hopes of identifying an enhancer. This method has the potential advantage that random screening might find activity the other more focused methods do not detect. However this method is out-dated, costly, more uncertain, and laborious.

The second method screens only non-coding sequences with high vertebrate homology. While homology does not necessarily equate to an enhancer, it is an important indicator as highly conserved regions have been found to be enhancer enriched (Bejerano et al., 2004; Pennacchio et al., 2006). Therefore, this method has a significantly better chance of identifying an enhancer than the previous method. In addition, a lot of sequencing data from numerous vertebrate species is freely and readily available.

Lastly, the third method screens areas where ChIP-seq has identified areas of histone modification and therefore possible sites with enhancer activity (Rada-Iglesias et al., 2012). This method relies on the temporal qualities of histone modification in a specific cell type, and can therefore target a specific moment and axial level in potential enhancer activity. If the Chip-seq data is not currently available, this would add additional time and cost if it needs to be obtained.

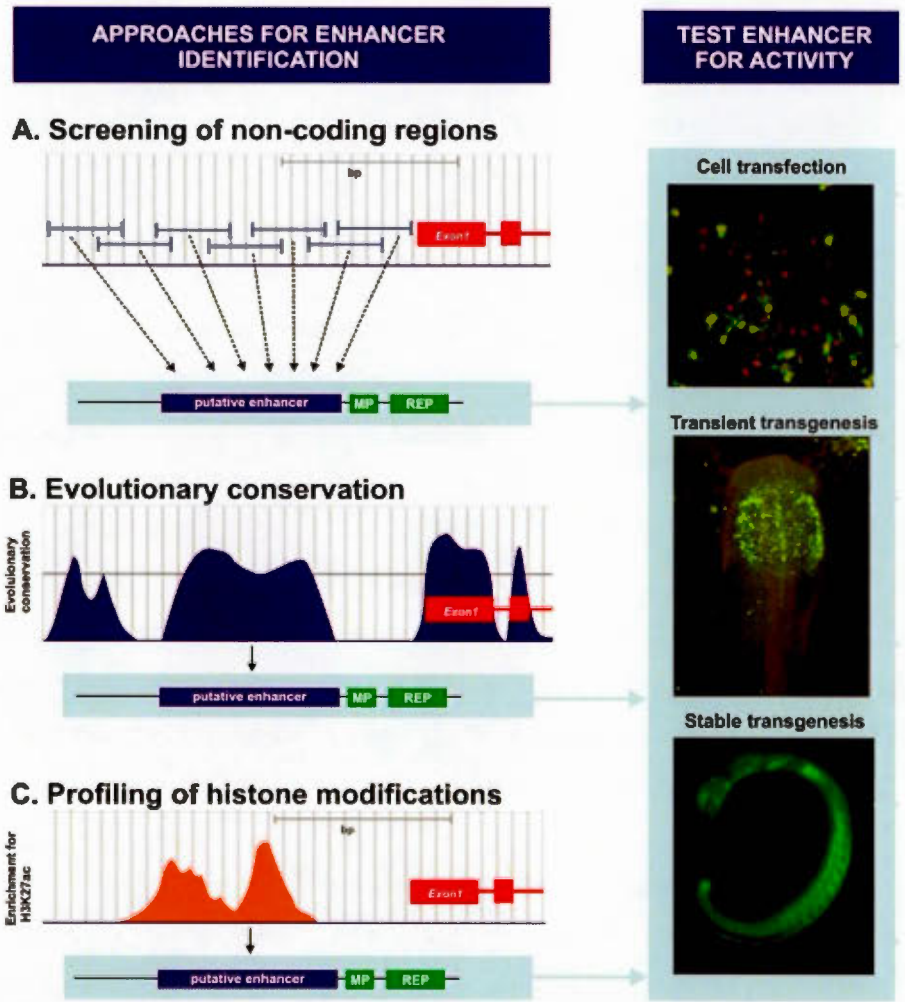


Figure 1.9 Techniques for identifying and testing enhancers.
First step: build a putative enhancer reporter construct, with the enhancer obtained from (A) Using a blind screen of non-coding areas; (B) Using a genome browser to identify conserved regions; (C) Using Chip-seq data to uncover areas of histone modification Second step: testing the enhancer by cell transfections, transient and/or stable transgenic models. Image retrieved from Simoes-Costa and Bronner (2013),

1.3.1.2 Characterization of binding motifs

While the previous methods can identify the existence of an enhancer, the location of activity, and even test several potential TF activators, the characterization of binding motifs requires different steps. Bioinformatic tools can be used to identify potential binding motifs but characterization is not addressed until the binding of transcription factor to DNA is demonstrated. Electrophoretic mobility shift assays (EMSA) are often used in promoter/enhancer studies to reveal the binding of transcription factors (Pares-Matos, 2013). A DNA pull-down assay followed by Mass-Spectrometry can also be used to identify DNA-binding TFs (Drewett et al., 2001; Hubner et al., 2015).

1.3.2 *Cdx2*NSE

The enhancer identification method, using evolutionary conserved regions, described above and in Figure1.9b, was used in the discovery of a neural specific enhancer (NSE) of *Cdx2*. Stable transgenesis showed reporter (LacZ) activity in the neural tube (Figure 1.10) (Wang and Shashikant, 2007).

1.3.2.1 Cre-LoxP application of *Cdx2*NSE

As well as providing evidence of regulatory interactions, enhancers can be used in a variety of applications, including fate mapping, time-lapse imaging and targeted loss of- and gain-of-function assays (Simoes-Costa and Bronner, 2013). One use of the *Cdx2*NSE is to exploit its neural specific activity and caudal localization in the developing vertebrate. Cre-LoxP mouse systems in particular, are a great tool and way of exploiting the temporal and spatial qualities of an enhancer. In this system, the enhancer drives expression of Cre-recombinase in a transgenic line of mice. This enzyme can recognize and excise DNA flanked by loxP sites. Conditional knockouts are generated from the successful cross of a tissue specific Cre-expressing line and a line carrying a floxed (loxP-flanked) gene of interest (Lodish et al., 2000).

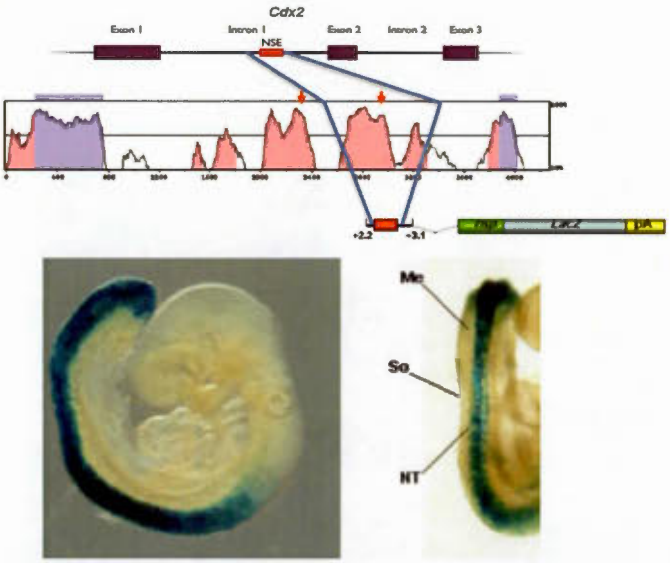


Figure 1.10 Identification and testing the Cdx2 neural specific enhancer (NSE). Cdx2NSE was identified by its high vertebrate homology, and tested by a transgenic line carrying the presumptive enhancer driving the activation of a LacZ reporter. Expression seems confined to the neural tube. Images retrieved from Wang and Shashikant (2007)

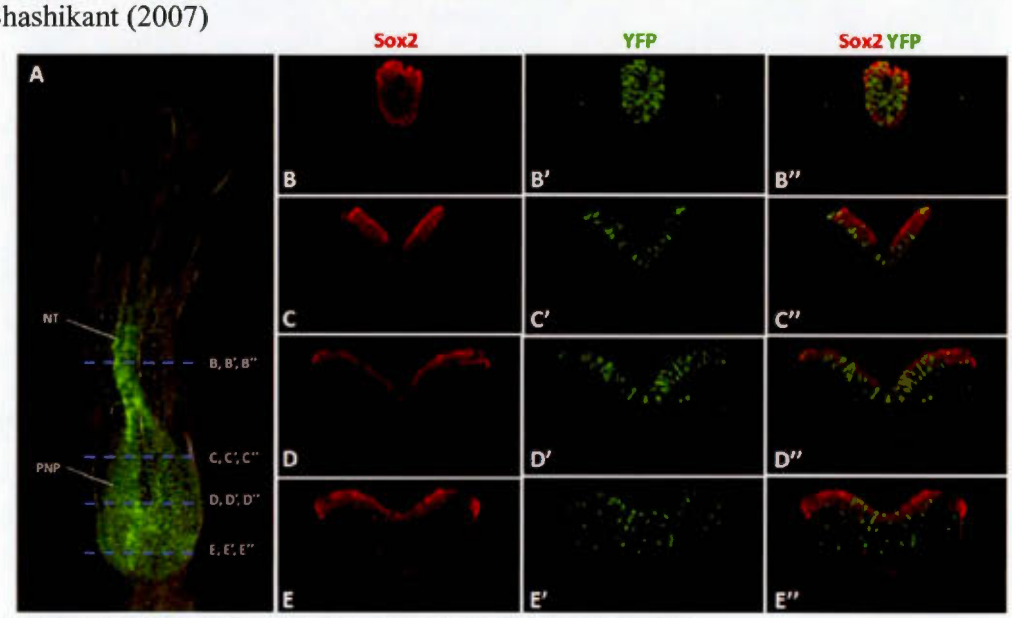


Figure 1.11 The enhancer is not entirely neural specific.
An e8.5 embryo obtained from a cross between *Cdx2*NSE-Cre driver line and R26R-YFP had Cre and therefore YFP expressed mostly in the neuroectoderm with some present in mesodermal cells. Compare YFP to neural marker Sox2. Image retrieved from Coutaud and Pilon (2013)

Here, the gene of interest is removed in specific Cre expressing tissues. Another use of the system is to have loxP sites impeding the expression of a gene of interest. In this example, the successful recombination leads to the tissue-specific expression of the target gene.

Seeing as how no such Cre-driver line targeting the posterior neuroectoderm existed, Coutaud and Pilon (2013) decided to generate such a line using the *Cdx2*NSE. However, once the Cre-driver line was created and tested it was revealed that the enhancer had minor activity in the mesoderm as well and was therefore not entirely neural specific (Figure 1.11).

1.4 Diseases of NCC origin and neural tube defects

A better understanding of the NC-GRN and associated epigenetic mechanisms is important because disruptions in these processes are believed to be the source of several human congenital birth defects and cancers. Relatedly, mechanisms affecting proper neurulation can lead to several neural tube defects (NTDs). The result of these regulatory disturbances vary, but may drive changes affecting cell-fate/programming decisions, differentiation timing and/or migration patterns.

1.4.1 Neurocristopathies

Neurocristopathies are NC-derived developmental anomalies comprising over 700* known syndromes and defects (Trainor, 2010). Some of the best-known birth defects due to NC malformation result in craniofacial defects, heart defects and aganglionosis of the colon. Many of these defects affect several known NC-GRN factors, such as Sox10, Pax3, Snail2 in Hirschsprung's disease and Waardenburg syndrome (Kim et

* For craniofacial defects alone

al., 2011; Paratore et al., 2002), or Sox9 and Twist in CHARGE syndrome (Hu et al., 2014).

1.4.1.1 Neural crest cell-derived cancers

NCCs possess many of the same attributes as metastatic cells, such as migration and invasion. The assumption then is that cells of NC origin restart many of the same regulatory mechanisms needed for embryonic development (Simoes-Costa and Bronner, 2013). Two of the most well-known cancers of NC lineage are neuroblastoma and melanoma.

Neuroblastoma arises from cell types of the sympatico-adrenal NCC lineage and is the most common solid childhood tumour accounting for 7-10% of paediatric cases (Brodeur, 2003; Jiang et al., 2011; Schulte et al., 2013). The change from NCC to mature sympathetic ganglia, involves an epigenetic switch needed to silence two genes that promote cell cycle progression and cell death aversion—respectively, *Mycn* and *Arid3b*—the same switch left on is believed to lead to the oncogenesis of neuroblastoma (Kobayashi et al., 2013). Mycn in particular has been shown to be expressed in early NCC during ventral migration and in NCC undergoing neuronal differentiation (Wakamatsu et al., 1997).

Studies have pointed out that the progression of metastatic melanoma may involve several genes of the NC specification module (of the NC-GRN) such as *Twist, Sox10, Slug* and *FoxD3* (Shakhova et al., 2012; Shirley et al., 2012). Slug is needed by NCC for EMT (Peinado et al., 2004), and Sox10 for early melanoblast formation and postnatal maintenance (Harris et al., 2013; Potterf et al., 2001). Twist1 promotes invasion of melanoma cells by mediating the regulation of a matrix metalloproteinase (MMP), an enzyme needed to breakdown the extracellular matrix (Weiss et al., 2012). *Twist* in particular is overexpressed in many solid tumours of NCC origin, not just melanoma, but also glioma and neuroblastoma, highlighting its importance in the metastatic process (Yang et al., 2004).

1.4.2 Neural tube defects

Neural tube defects are congenital deformities brought about by an inadequate closure of the neural tube. Prevalence of NTDs is from 0.75 to 1.12 per thousand, depending on ethnic origin (Feuchtbaum et al., 1999). NTDs are complex in that a diet rich in folic acid may well be an important factor in prevention, as public health data seems to indicate (Czeizel et al., 2011; Erickson, 2002; Pitkin, 2007). Interestingly, there is also evidence indicating that the NCC may also be protected by folate supplementation, suggesting the neuroepithelium may be sensitive to folate-related pathways (Antony and Hansen, 2000; Li et al., 2011). Other possible NTD risk factors may include environmental exposure to teratogenic agents such as heavy metals, organic solvents and agricultural chemicals (Sever, 1995). Depending on where the anomaly occurs the result is usually *spina bifida* or anencephaly, although there are several others. Improper closure near the caudal neuropore leads to *spina bifida* and anencephaly in the rostral neuropore. The degree of affliction varies, as does neurological dysfunction. In *spina bifida* cognition, behavior, vertebral column and organ systems can be affected depending on severity (Fletcher and Brei, 2010). Anencephaly is usually more severe, because the neural folds of the brain remain open (Copp and Greene, 2013) leaving a large portion of the brain undeveloped. Survival rate is very low, i.e. 40% after 24 hours (Baird and Sadovnick, 1984). Proper closure of the cranial and spinal neural tube depends a lot on the bending of the neuroepithelium at median (MHP) and dorsal-lateral hinge points (DLHPs). MHP and DLHPs are both features of cranial and intermediate spinal neurulation. However, MHP is a feature of upper spinal neurulation and DLHP of the lower spine (Copp et al., 2003; Yamaguchi and Miura, 2013; Ybot-Gonzalez et al., 2007). BMP antagonist Noggin is needed for cranial as well as lower and intermediate spinal DLHPs (Stottmann et al., 2006). Conversely, the upper spine MHP relies on inhibition of Noggin by Sonic-hedgehog, hence higher BMP levels (Ybot-Gonzalez et al., 2002; Ybot-Gonzalez et al., 2007).

1.4.3 Tools for studying neurocristopathies and NTDs

Many of the molecular mechanisms so far described have been clarified through the use of mutant mouse models (Britsch et al., 2001; Dixon et al., 2006; Shakhova et al., 2012; Southard-Smith et al., 1998; Weiss et al., 1997). Dom (Dominant megacolon) mice are mutant (chromosome 15) models for Hirschsprung's disease that arose spontaneously, i.e. were not purposefully engineered, and they were instrumental in demonstrating the importance of Sox10 for development of the peripheral/enteric nervous system (Puliti et al., 1995; Southard-Smith et al., 1998). *Splotch* mice also arose spontaneously (Russel, 1947), with a mutation in chromosome 1 (Epstein et al., 1991a) and are often used as models for NTDs (Moase and Trasler, 1992). In homozygous mice, the mutation is embryonic lethal by E14 (Li et al., 1999), exhibiting NT and heart defects (Auerbach, 1954; Conway et al., 1997). The heterozygous *Splotch* line carries one null *Pax3* allele, and is viable having minor pigmentation abnormalities. In the early 90's when the *Pax3* deletion was found to be the cause of the mutation, it quickly indicated the importance of Pax3 for normal neural development (Epstein et al., 1991b; Goulding et al., 1993).

Conditional knockouts (KO) have also been used, many exploiting the NC-specific *Wnt1*-Cre transgenic line (Danielian et al., 1998). Wnt1 is active in the embryonic NT and migrating NCC, and is important for the formation of the midbrain–hindbrain boundary (Echelard et al., 1994; Lewis et al., 2013). Using this NC/NT specific Cre-driver line, Dudas et al. (2004) knocked-out a BMP receptor (Alk2) which then produced craniofacial defects including cleft palate and mandible malformation and suggested the importance of Alk2 in regulating the formation of specific cranial features. Akiyama et al. (2004) employed it to generate conditional *Sox9* KO embryos to examine cardiac NCC and the essential role of Sox9 in heart EMT. Degenhardt et al. (2010) used it to rescue lost *Pax3* expression specifically in the NC, and demonstrated the redundancy of *Pax3* "neural crest enhancers". Other Cre-driver

lines used include *Nestin*-Cre for the neuroectoderm (Dubois et al., 2006; Tronche et al., 1999) and *Pax3pro*-Cre for the dorsal NT and NC (Li et al., 2000).

Overexpression knock-in mutants are also instrumental models for disease. Gene amplification in particular is known to be the direct cause of many known human diseases and cancers (Santarius et al., 2010; Shastry, 1995). For example, the classic mainstay mouse model for neuroblastoma is the TH-MYCN, wherein *Mycn* expression is driven by the TH rat tyrosine hydroxylase promoter (Weiss et al., 1997). As previously discussed, *Mycn* is normally expressed in migrating NCC switching to neuronal lineage (Wakamatsu et al., 1997). At the time, *MYCN* was known to be a proto-oncogene of human neuroblastoma but had not yet been tested in an animal model. Weiss et al. (1997) provided direct *in vivo* evidence that *Mycn* amplification can contribute to the oncogenesis of neuroblastoma. The TH-MYCN transgene is integrated in chromosome 18 in a distal region, and the effect it has in disturbing this region is unknown. Due to this uncertainty, Althoff et al. (2015) created another *Mycn*-driven neuroblastoma mouse model targeting the *ROSA26* locus (chromosome 6). This locus has the advantage that it is ubiquitously expressed, disruptions in the locus have no physical effect in mice, and can be targeted with high efficiency (Soriano, 1999; Zambrowicz et al., 1997). As of 2010, there were already over 130 knock-in lines targeting the *ROSA26* locus (Casola, 2010).

All these mutant models have their own limitations, and advantages, and yet they share the fact that the disruptions they create help fuel our understanding of the molecular mechanisms involved in the development of NC and NT anomalies.

CHAPTER II

HYPOTHESES AND OBJECTIVES

2.1 Hypotheses

Conditional overexpression of *Cdx1* plays an oncogenic role in the development of neural crest cell derived cancers in envisioned mouse model.

The intronic neural specific enhancer (NSE) of *Cdx2* contains identifiable non-neural-specific activator binding sites.

2.2 Objectives

Generate a mouse model carrying the Cre-inducible Cdx1 overexpression transgene. The steps required to generate the model: Construction of the targeting vector, transfection of the vector into ES cells, selection of the colonies, verification of positive clones, microinjection and rearing.

Investigate regulation of *Cdx2*NSE, and identify activator-binding sites with non-neural specific characteristics. Testing the regulation would give us a better understanding as to how the NSE architecture could be modified to drive truly neural specific expression in a new Cre-driver line. Luciferase assays using neural Sox2 could identify which areas are potentially neural specific.

CHAPTER III

MATERIALS AND METHODS

3.1 Construction of pROSA26-Cdx1, ES targeting and verification

3.1.1 Vector construction

To make the pROSA26-Cdx1 targeting vector, a PGKneoPA-Flag-Cdx1-IRES-EGFP-BGH cassette was inserted into the PacI and AscI sites of the pROSA26-PA targeting vector (Srinivas et al., 2001). pROSA26-PA was a gift from Frank Costantini (Addgene plasmid # 21271). It is a vector used to target the ubiquitous *ROSA26* locus, and contains the 3' and 5' *ROSA26* arms necessary for homologous recombination (see Figure 4.2) as well as a diphtheria toxin gene (PGK-DTA) for negative selection in ES cells. The final construct was digestion and sequencing verified. In total, the final construct contains a loxP-flanked PGKneoPA cassette capable of kanamycin resistance expression in *E.coli* and eukaryotic promoter PGK for expression of neomycin resistance in the mammalian ES cells. Next the N-terminal FLAG-tag is upstream of a full *Cdx1* ORF (807-bp) as previously described (Beland et al., 2004), followed by an IRES-EGFP reporter sequence.

3.1.2 Targeting ES cells

R1 ES cells were cultured on mouse embryo fibroblast (MEF) feeder cells that were mitotically inactivated by 10μg/ml mitomycin C treatment. The culturing of the ES cells on MEF feeder cells was done on gelatin-coated plates to improve conditions. Also, the ES medium was supplemented with leukemia inhibitory factor (LIF) to reduce possible differentiation of ES cells. The pROSA26-Cdx1 vector was

linearized by SacII and (12.5 µg DNA) electroporated (250V, 500 µF, in a 4mm cuvette) into the R1 ES cells. Selection of stable integrants was done with 200µg/ml G418 for one week. Genomic DNA was extracted and isolated. Screening of successful homologous recombination was carried out first by PCR using an external forward ROSA26 primer and an internal reverse primer to create a 1.1kb fragment (Table 3.1). Three clones were identified. Confirmation by Southern blot, albeit inconclusive, was followed soon after.

Table 3.1 Primers for PCR screen.

PRIMER NAME	LENGTH	SEQUENCE
ROSA26TARG-F1	23	AAGAAGAGGCTGTGCTTTGGGGC
ROSA26TARG-R1	20	AGGGCGGCTTGGTGCGTTTG

3.1.3 Southern Blotting

Around 20 µg of genomic DNA was digested with EcoRI and KpnI and separated on a 0.8% agarose gel. The gel was then depurinated, denatured and neutralized. Afterwards, the DNA was transferred over night onto a Hybond membrane (Amersham) via a simplified downward capillary system, i.e. without added transfer buffer.

Once transferred, the membrane was rinsed with 2x SSC, dried, UV cross-linked and pre-hybridized. The pre-hybridization buffer solution contained salmon sperm ssDNA to block non-specific sites. After pre-hybridization, the membrane was hybridized with probe ssDNA.

Two DNA probes were prepared to target the 5' or 3' *ROSA26* ends (Figure 4.2), at 25ng in 45µl of TE buffer. The probes were [α-^{32}P] dCTP-labelled using the Rediprime II labelling kit (Amersham), and purified by sepharose columns. The final probe concentration was approximately around 2.8ng/mL of hybridization buffer, as per manufacturer's instructions. Hybridization occurred over night at 65°C. The

membrane was later washed at least 2 times (2x SSC + 0.1% SDS) before placing in an exposure cassette for 1-7 days.

3.2 Testing the neural specific regulation of *Cdx2*NSE

3.2.1 Plasmids

The full *Cdx2*NSEforward-LUC and *Cdx2*NSEreverse-LUC reporter constructs were created in our lab (Coutaud, 2013) and were readily available. They contain an 852-bp *Cdx2*NSE sequence upstream of an *Hsp*68 minimal promoter in a luciferase gene expression vector. Four *Cdx2*NSEreverse fragments were created by PCR with the forward primers containing a HindIII site and reverse primers containing a KpnI cut sites (Table 3.3). PCR products were amplified with Taq polymerase (Feldan) and cloned into a pGEMT-easy vector (Promega) and verified by sequencing. The HindIII and KpnI restriction sites were then used to clone into an *Hsp*68-LUC reporter vector. The resulting four fragment NSE reporter constructs can be seen in Figure 5.4, and include the NSE sequences containing putative Lef1/Tcf sites : 4-3-2-1 (676-bp); 3-2-1 (607-bp); 2-1 (437-bp); and 5 (193-bp). The construct with LEF/TCF site « 1 » (219-bp) was made by digesting *Cdx2*NSEreverse-LUC with BamHI to remove a 666bp fragment containing potential Lef1/Tcf binding sites 5-4-3-2. PGL3-OT served as positive control in transfections; it contains 3 copies of wildtype Tcf-4 binding sites and was a gift from Bert Vogelstein (Addgene plasmid # 16558) (Shih et al., 2000). Negative control was an empty vector containing just the promoter and luciferase gene. Serving as the activators for the transfections were Sox2, and Lef1-βcatenin.

3.2.2 Luciferase assay conditions

P19 cells were cultured at 37°C (5% CO_2) in α-MEM supplemented with 7.5% heat-inactivated bovine calf serum and 2.5% heat-inactivated fetal bovine serum (Wisent). Approximately 2 hours prior to transfection, the P19 cells were plated in 24 well plates with a cell density of $3x10^4$ cells/well and incubated at 37°C (5% CO_2).

Co-transfection was carried out to test the effect of Sox2 and Lef1-βcatenin (individually and together) on the full *Cdx2*NSE forward and reverse reporter sequences. *Cdx2*NSEforward-Luc and *Cdx2*NSEreverse-Luc reporter DNA was maintained at 100ng/well, as was the negative and positive control DNA. On the other hand, Sox2 and Lef1-βcatenin activator DNA ranged from 0ng, 25ng, 50ng or 100ng/well. The total amount of DNA per well was kept at 300ng. If needed, an empty IRES-GFP expression vector was used to complete this amount. GeneJuice was the transfection reagent (Novagen) and the ratio of reagent to DNA was 3µl to 1µg DNA, as recommended, i.e. each well had 0.9µl of transfection reagent and 300ng of total DNA. This set of transfections was performed at least three times in triplicate.

Assay conditions were similar for the tests with the truncated *Cdx2*NSEreverse DNA reporter sequences. However, this set underwent two key differences in conditions: the cell density used and activator DNA concentration. Cell density was kept at $1.5x10^4$ cells/well and activator DNA (Sox2 and/or Lef1-βcatenin) was kept at 0ng or 100ng/well. This set of transfections was performed at least six times in triplicate.

All transfections were incubated in the α-MEM FBS+CBS media for about 48 hours at 37°C before performing post-transfection tests and analyses. After 48 hours, transfection efficiency was visually assessed by microscopy, i.e. identifying GFP expression. Next, transfection analysis (gene expression) was quantified by luminometry. The steps for this included: media removal, rinsing wells with phosphate buffered saline 1X, and lysing the cells. In a tube, 20µl of the cell lysate was added to 100µl of luciferase assay buffer. Then, 50µl of luciferin was delivered before placing tubes in the luminometer,

Luciferase activity was expressed as fold activation relative to the appropriate reporter vector alone. Each independent experiment was carried out in triplicate, at least three times.

3.2.3 Luciferase assay test and statistical analysis

All the results of the luciferase assays are expressed as means+S.D., and the differences in luciferase activity are expressed as fold activation. The differences in fold activation were examined by student's t-test for two-group comparisons. One-way ANOVA was performed to identify differences in multiple group comparisons Graphs and all statistical analyses were done using GraphPad Prism version 6 software, and p-value ranges, wording and asterisk representation follow the GraphPad statistics guide (Table 3.2).

Table 3.2 P-value guide

P VALUE	DESCRIPTION	REPRESENTATION
< 0.0001	Extremely significant	****
0.0001 to 0.001	Extremely significant	***
0.001 to 0.01	Very significant	**
0.01 to 0.05	Significant	*
≥ 0.05	Not significant	NS

Table 3.3 Oligos to construct *Cdx2*NSE reverse strand fragments.

PRIMER NAME	LENGTH	SEQUENCE (5' to 3')
Cdx2NSE-antisense 4-3-2-1 LEF1/TCF-F	31	AAGCTTACAGAATGCTGGCCAGGAACTGTTC
Cdx2NSE-antisense 3-2-1 LEF1/TCF-F	35	AAGCTTTAGTGCTCAGTGCCTAGTTGAACAAC CAG
Cdx2NSE-antisense 2-1 LEF1/TCF-F	29	AAGCTTGGGAACACAGTCGCAAACAATGC
Cdx2NSE-antisense-R	30	GGTACCGTGCTCTAAGAGCAGCATCCGTTC
Cdx2NSE-antisense 5 LEF1/TCF-F	35	AAGCTTCTATCCTGACCAAGTGACCTGTGATC ATT
Cdx2NSE-antisense-R	26	GGTACCGAACAGTTCCTGGCCAGCAT

CHAPTER IV

TOWARD THE GENERATION OF MICE CARRYING THE CRE-INDUCIBLE *CDX1* TRANSGENE

4.1 Introduction

Disruptions in NCC development can cause several neurocristopathies and cancers, including craniofacial defects, heart defects, colon aganglionosis, neuroblastoma, and melanoma among others (Hall, 1999; Trainor, 2013). The aim of this project was to generate a knock-in mouse capable of conditional *Cdx1* overexpression. Considering the key role Cdx proteins play in NC development, conditional overexpression of this key NC-GRN player could engender the pathogenesis of a NC derived cancer. NCC derived tumours, such as melanoma and neuroblastoma, exhibit overexpression of genes needed for NC development, in particular those affecting EMT and migration (Trainor, 2013). It is the hope that this tool will help identify previously unknown regulatory components linking the contributions of Cdx neural functions to malignancy.

4.2 Results

4.2.1 Construction of the targeting vector

The ROSA26 targeting vector was chosen to knock-in the transgene into the mouse genome in the *ROSA26* locus (mouse chromosome 6) (Soriano, 1999; Srinivas et al., 2001). Targeting this locus ensures the transgene will be stably integrated in one place and will be ubiquitously expressed by an endogenous *ROSA26* promoter. In addition, disruptions to the locus do not affect mouse viability or cell phenotype.

With this in mind, we used a vector containing the bicistronic FLAGCdx1-IRES-GFP cassette that had previously been cloned into a BigT vector. In this vector, a floxed-PGK-neo-tpA cassette functions as a stop cassette for downstream FLAGCdx1-IRES-GFP. The fragment containing loxP-PGK-neo-tpA-loxP-FLAGCdx1-IRES-GFP-bpA was then cloned into the PacI and AscI restriction sites of pROSA26PA vector (Figure 4.1) (Srinivas et al., 2001). The final 15-kb construct was digestion and sequence verified. Once confirmed, 12.5 μg of SacII-linearized targeting construct DNA was electroporated into R1 ES cells.

4.2.2 Targeting the *ROSA26* locus

After electroporation, the cells were expanded in medium containing the selection agent G418 for a week. Genomic DNA from 96 colonies resistant to G418 was extracted and analyzed by PCR (Figure 4.3) and Southern blot (not shown). Figure 4.2 demonstrates the entire targeting process. The targeting vector pROSA26-Cdx1 linearized and transfected targets the mouse ES cell *ROSA26* locus using the *ROSA26* 5' and 3' flanking arms. Successful recombination was screened by F1 and R1 PCR primers; three successful events were detected out of 96 colonies (Figure 4.3). The next step was to use Southern blot to further verify successful transgene targeting. Here the genomic DNA of the 3 PCR-verified clones was digested by EcoRI and KpnI, and the use of the 5' and 3' probes would detect a 4-kb an 8.8-kb fragment where the targeting event was successful, and in 11-kb fragment in the wild-type. Four successive Southern blots were carried out yet failed to even detect the control (WT).

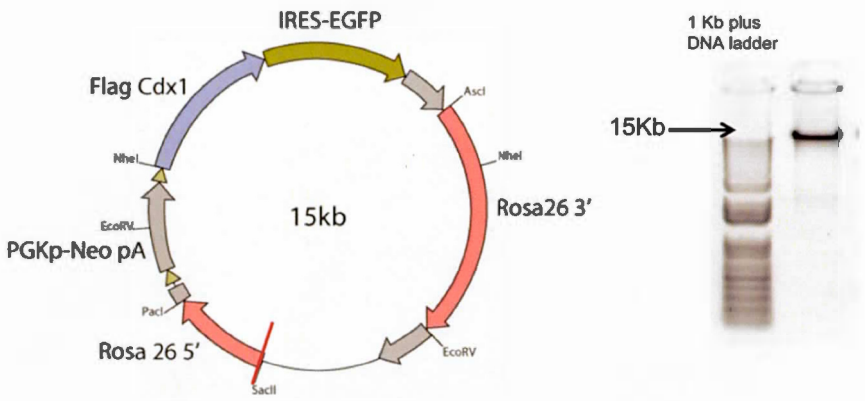


Figure 4.1 SacII linearization of pROSA26-Cdx1
Linearization of the pROSA26-Cdx1 targeting vector by SacII

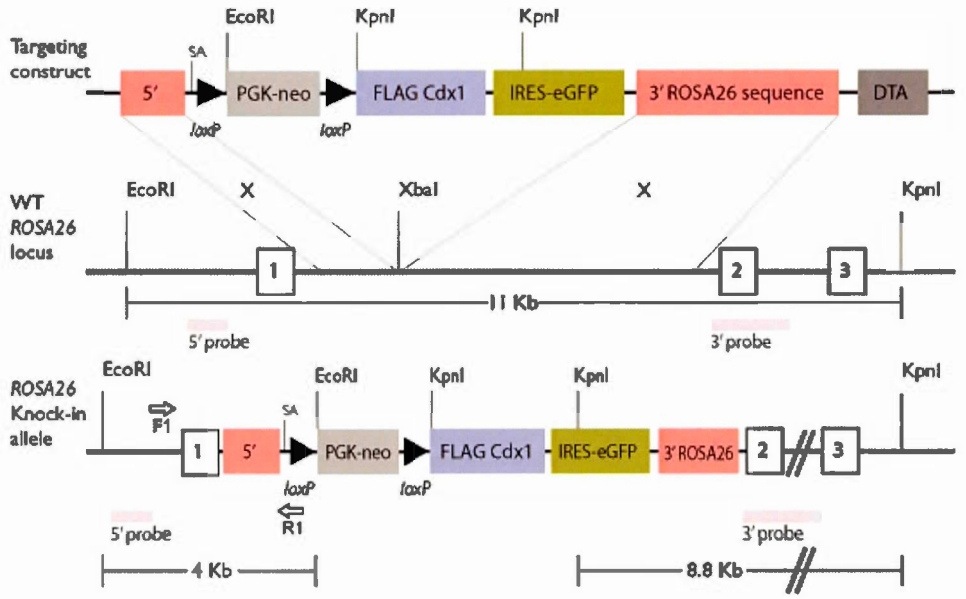


Figure 4.2 Depiction of gene targeting at the *ROSA26* locus.
At the top is the linearized targeting construct pROSA26-Cdx1. In the middle is the wild-type *ROSA26* locus, with exons 1, 2 and 3. At the bottom is the targeted knock-in allele containing a floxed-PGK-neo-tpA cassette which functions as a stop cassette, interrupting the expression of a bicistronic cassette containing FLAG-tagged Cdx1 and IRES-eGFP. PCR screening of successful knock-ins by the external F1 and internal R1 primers give a 1.1-kb product. Southern blot screening uses external *ROSA* 5' and 3' probes. ES cell genomic DNA digestion by EcoRI and KpnI permits the 5' probe to bind to a 4-kb fragment and the 3' to an 8.8-kb fragment where the targeting event is successful; in the WT allele the 2 probes would bind to an 11-kb fragment.

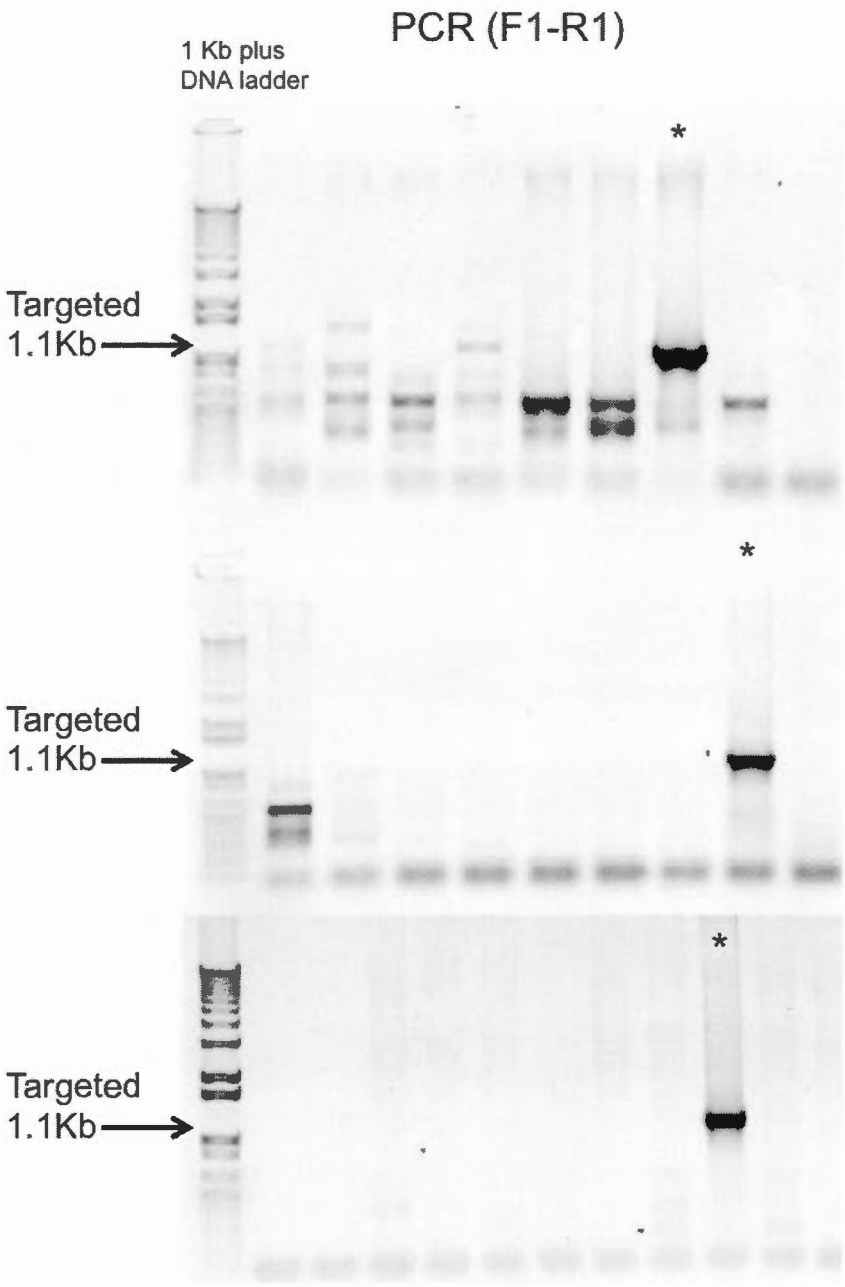


Figure 4.3 PCR screen of targeted ES clones.
PCR screen of 96 clone DNA samples yielded 3 correctly targeted events (marked with an asterisk).

CHAPTER V

TESTING THE NEURAL SPECIFIC REGULATION OF *CDX2*NSE

5.1 Introduction

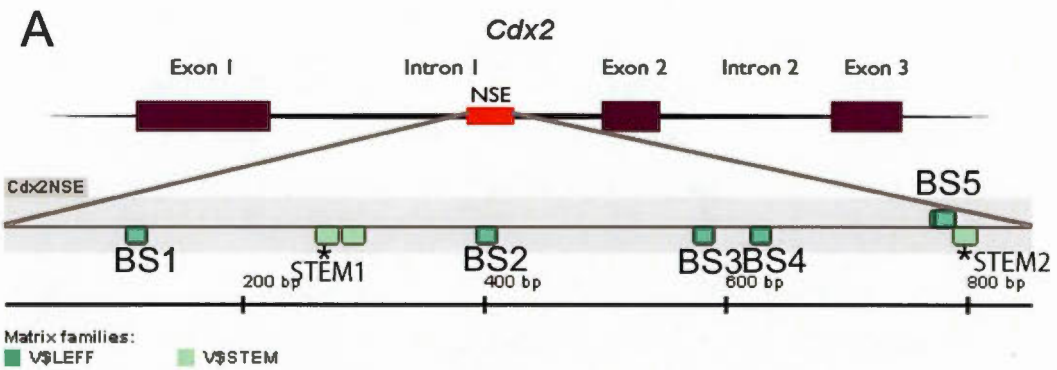
As previously discussed, *Cdx* genes integrate several of the posteriorizing signals important for development including Wnt, FGF, and RA signals (Bel-Vialar et al., 2002; Houle et al., 2003b; Keenan et al., 2006; Lohnes, 2003; Pilon et al., 2006; Pilon et al., 2007; Shimizu et al., 2005). Canonical Wnt signalling in particular is important for inducing *Pax3* expression in the posterior NPB, and with Cdx proteins as intermediaries (Sanchez-Ferras et al., 2014; Sanchez-Ferras et al., 2012). For this reason, as well as evidence demonstrated by others (Pilon et al., 2006; Shimizu et al., 2005; Zhao et al., 2014), *Cdx* genes can be said to be prime targets of Wnt signalling. In this project, the nuclear effectors of the canonical Wnt pathway, Lef1-βcatenin, were used in the transfections aiming to simulate the pathway and activate the neural specific enhancer (NSE) of *Cdx2*. Since the goal is to identify and distinguish the sites responsive to neural specific activation (and in its absence the potentially non-neural ones), it was necessary to add Sox2—a neural specific transcription factor and early marker of the neural plate (Papanayotou et al., 2008; Pevny and Nicolis, 2010; Rex et al., 1997).

5.2 Results

5.2.1 Potential binding activators

MatInspector is an online software used to predict the potential transcription factor binding site (TFBS) location on a query DNA sequence, that is, a promoter or enhancer. The potential TFBSs detected when submitting the 852-bp NSE sequence to the genomatix servers was impressive, at 256, with matrix similarity set to the optimized >0.75; greater than 0.80 is classified as a "good" match. Mindful of the importance of Wnt signalling, potential TFBSs of the nuclear effectors Lef1/Tcf were searched. There were 6 potential sites (Figure 5.1a,b). One site overlapped with another and was therefore counted as one, and will hereafter be referred to as one; it was also the only potential TFBS for Lef1/Tcf effectors found in the forward NSE strand. Matrix similarity ranged from 0.84-0.96, the highest similarity scores came from the reverse NSE strand.

Sox2 TFBS were found only as part of a composed binding site, which includes Oct4, Nanog, Tcf3 (Tcf7l1) and Sall4b (Figure 5.1a,c), and classified as STEM by the software. There were only two of these composed binding sites, and matrix similarity was marginally below the cut-off range for "good" matrix similarity of >0.80. These sites were located on the reverse strand exclusively.



B LEF1/TCF

Matrix Family	Detailed Family Information	Matrix	Detailed Matrix Information	Position from	Position to	Position anchor	Strand	Matrix sim.	Sequence (red: ci-value > 60 CAPITALS: core sequence)
V$LEFF	LEF1/TCF	V$LEF1.01	TCF/LEF-1, involved in the Wnt signal transduction pathway	103	119	111	(-)	0.922	cgaatagCAAAgagctt
V$LEFF	LEF1/TCF	V$LEF1.02	TCF/LEF-1, involved in the Wnt signal transduction pathway	393	409	401	(-)	0.962	atcttatCAAAgaaaag
V$LEFF	LEF1/TCF	V$LEF1.03	TCF/LEF-1	573	589	581	(-)	0.930	aggagaTCTAaggggtg
V$LEFF	LEF1/TCF	V$TCF7L1.01	HMG box transcription factor Tcf7l1 (TCF3)	620	636	628	(-)	0.888	ctcacatCAAAacagat
V$LEFF	LEF1/TCF	V$LEF1.04	TCF/LEF-1 (secondary DNA binding preference)	769	785	777	(+)	0.847	agtaaaACAAtcaatcc
V$LEFF	LEF1/TCF	V$LEF1.04	TCF/LEF-1 (secondary DNA binding preference)	773	789	781	(+)	0.895	aaacaaTCAAtccaaga

C STEM

Matrix Family	Detailed Family Information	Matrix	Detailed Matrix Information	Position from	Position to	Position anchor	Strand	Matrix sim.	Sequence (red: ci-value > 60 CAPITALS: core sequence)
V$STEM	Motif composed of binding sites for pluripotency or stem cell factors	V$OSNT.01	Composed binding site for Oct4, Sox2, Nanog, Tcf3 (Tcf7l1) and Sall4b in pluripotent cells	260	278	269	(-)	0.789	cgcaggcGCATttgaatgg
V$STEM	Motif composed of binding sites for pluripotency or stem cell factors	V$OSNT.01	Composed binding site for Oct4, Sox2, Nanog, Tcf3 (Tcf7l1) and Sall4b in pluripotent cells	788	806	797	(-)	0.782	tgagtcaACATattaaatc

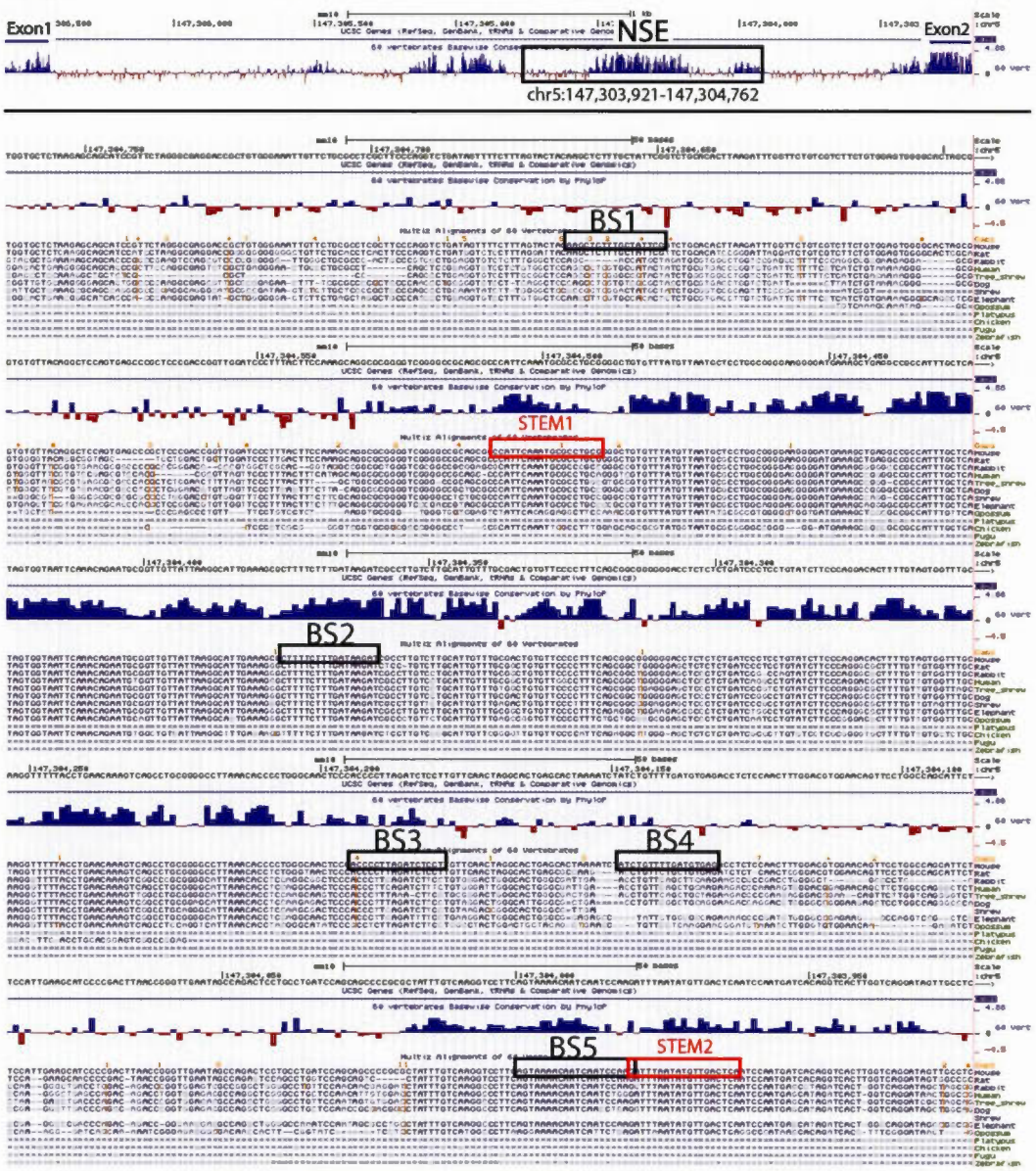
Figure 5.1 Potential Lef11/Tcf and STEM family binding sites.
(A) Forward and reverse strand locations of potential TFBS for (B) the canonical Wnt pathway effectors of the Lef1/Tcf family (five total Binding-Sites, BS; shown in dark green), and (C) the combined binding sites of Oct4, Sox2, Nanog, Tcf3 (Tcf7I1) and Sall4b in pluripotent cells (two total STEM sites; shown in light green). The third STEM site seen in schematic A is an Oct3 and not a Sox2 binding site, and is thus excluded. BS6 is also excluded because it overlaps with BS5. Instead, for practicality, BS5 combines both. Core sequences are the four most conserved consecutive nucleotides. Higher consensus index values indicate greater conservation. Images retrieved from MatInspector.

5.2.2 NSE and potential TFBS sequence homology

The next step in analyzing the potential TFBSs is observing how they match up with vertebrate conserved regions. A conserved TFBS may indicate evolutionary selection and therefore a good candidate for further testing. The online University of California Santa Cruz (UCSC) Genome Browser allows us to visually discern how conserved the domains are. In Figure 5.2a, we can see the entire enhancer, in the forward direction, and how some potential binding sites are more conserved than others. The phyloP scores of 60 vertebrates, which include 40 placental mammals, determine basewise conservation. The higher the positive score is, the greater the conservation. Negative scores indicate faster evolution (Rhead et al., 2010). The range for the entire enhancer was from -5.114 to 4.8243.

Figure 5.2b, is an expanded view of the binding sites found in Figure 5.2a. Potential Lef1/Tcf binding-site-1 (BS1) has the least sequence homology, with a mean phyloP score of -0.381687. BS2 has the highest conserved sequence, with a mean phyloP score of 3.65525. Both conserved "STEM" binding sites are also highly conserved. See Appendix Table A.1 for phyloP scores of each potential TFBS sequence as well as of their highest conserved core sequence.

A


Exon1
NSE
Exon2
chr5:147,303,921-147,304,762
BS1
STEM1
BS2
BS3
BS4
BS5
STEM2

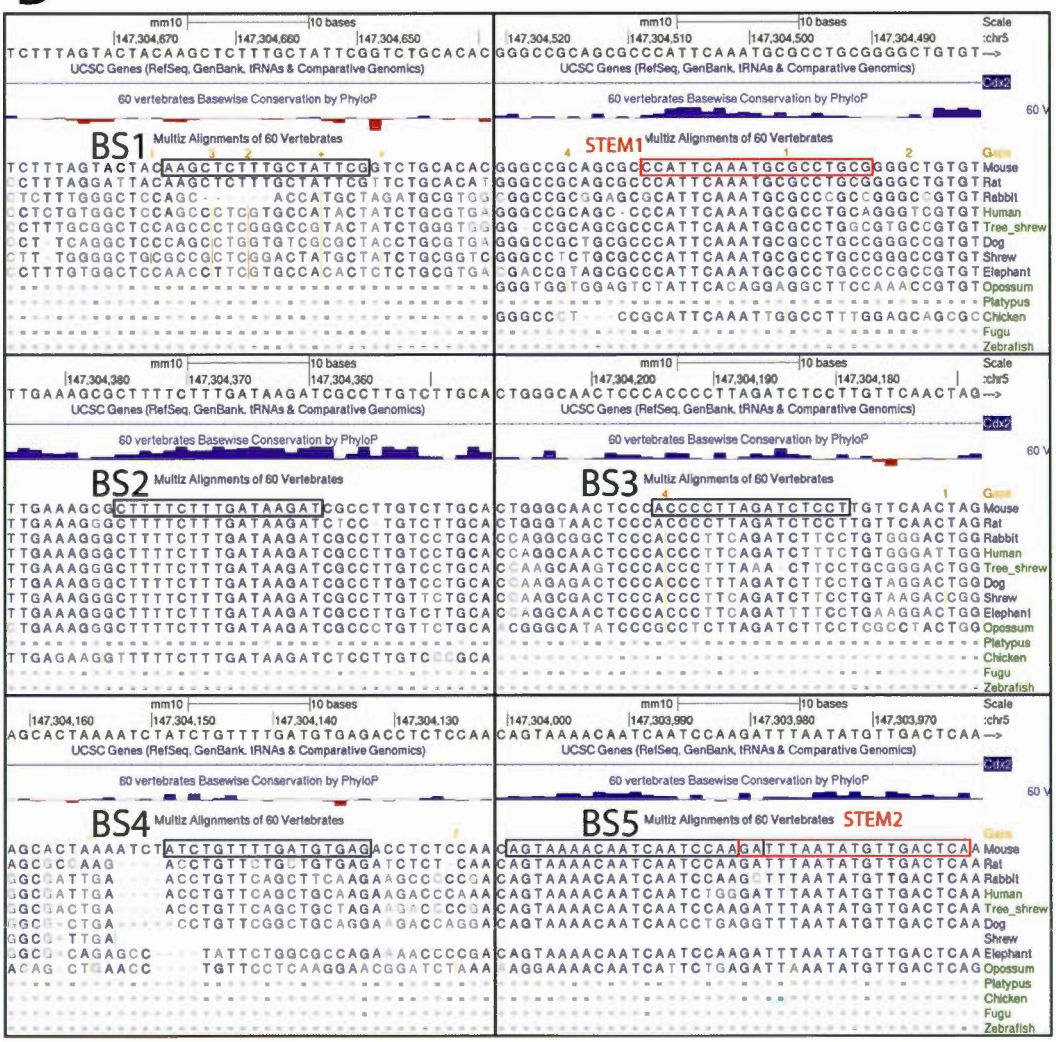


Figure 5.2 Vertebrate homology of the neural specific enhancer.
Outlined in black are the potential TFBS for the Lef1/Tcf family of effectors (Binding-Sites 1-5), and in red, the composed binding sites for Sox2 and other pluripotent/stem cell markers (STEM 1/2). Panel B is an expanded view of Panel A. The basewise conservation of 60 vertebrates is plotted as positive (conserved) or negative (acceleration) phyloP scores. Out of the 60 vertebrates, 40 are placental mammals and, of these, 8 are rodents, pika and rabbit. For clarity, 13 of the 60 vertebrates are represented in the sequence alignment. Taken from the UCSC genome browser.

5.2.3 Directional preference and synergy

Luciferase assays were carried out to test the potential activators Sox2 and Lef1-βcatenin on the full *Cdx2*NSE forward and reverse reporter sequences. Transfections were carried out on P19 cells, which are pluripotent embryonic carcinoma that have the ability to differentiate into muscle and neuron cell-types (McBurney, 1993) Lef1-βcatenin are the nuclear effectors of posteriorizing Wnt signalling, and Sox2 is a neural specific transcription factor. The potential activators were either transfected alone, or co-transfected together in doses that ranged from 0-100ng DNA each with 100ng of the Cdx2NSE luciferase reporter in either reverse or forward NSE strand. The reverse strand displayed significant (2.2 times the sum of each) synergistic activation with both Lef1-βcatenin and Sox2. Also, the reverse strand experienced very significant activation ($p = 0.006$) compared to the forward strand in co-transfection (Figure 5.3). This finding was surprising considering how enhancers do not generally display directional preference (Pennacchio et al., 2013). Interestingly, all potential TFBS except BS5 are located on the reverse strand. This test also determined the dose (100ng) needed for the displayed synergy.

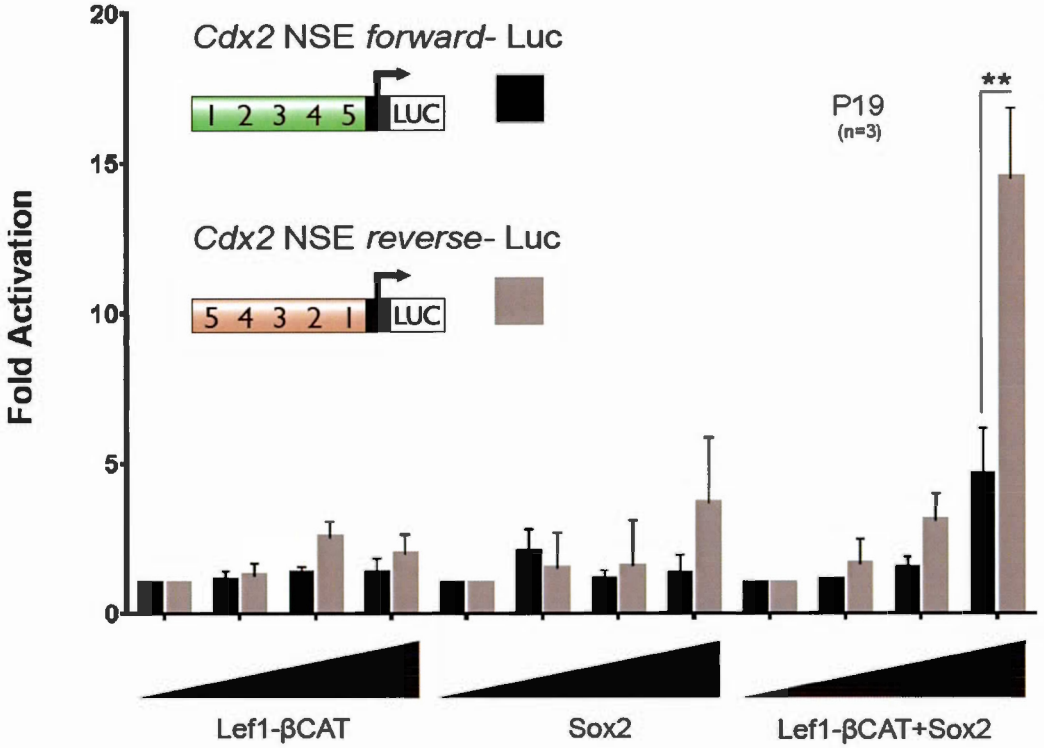


Figure 5.3 *Cdx2*NSE synergy and directional preference.
Luciferase assay tested the enhancer in the forward and reversed direction. Co-transfections utilized Lef1-βcatenin to simulate the effectors of posteriorizing Wnt, and/or neural specific Sox2. Doses ranged from 0-100ng of activator DNA with 100ng of *Cdx2*NSE. The reverse strand displayed significant synergy with both Lef1-βcatenin and Sox2 (**) indicates $p = 0.006$ (t-test). Luciferase activity for each condition was calculated as fold activation compared to the reporter vector alone. The results are represented as the average + standard deviation of triplicate samples from three experiments.

5.2.4 Testing the reverse NSE

Since the reverse NSE was identified as the preferred direction for Lef1-βcatenin and Sox2 activation, it merited additional transcriptional activation analysis. Fragmenting the reverse NSE in smaller units permits us to test a narrower group of binding domains. The reverse NSE was fragmented five times containing potential Lef1/Tcf binding sites: (BS4,3,2,1), (BS4,3,2,1), (BS3,2,1), (BS2,1), (BS1), and (BS5). The location of the STEM (Sox2) composed binding sites was incidental, that is fragmenting was focused on the potential binding sites for Lef1/Tcf. Removing BS5 in particular was important considering the binding motif is located on the forward NSE strand. Surprisingly, compared to the full reverse NSE strand, the strand missing BS5 had on average 0.7x the activation of the full sequence (Figure 5.4). Yet BS5 by itself did not display any activation. There was also a significant discrepancy between the activation seen here for the full reverse NSE strand and what was seen previously (Figure 5.3); possible reasons for this will be discussed in the following chapter. Lef1-βcatenin and Sox2 synergy was only apparent in the full reverse strand.

In summary, none of the fragments recapitulated the neural specific activation or synergy found in the full reverse strand as hoped. Thus it remains unknown which areas of the NSE architecture are clearly sensitive to neural specific activation and which areas are not. In addition, it is unclear if the STEM (Sox2, Oct4, Nanog, Tcf3/ Tcf7l1 and Sall4b) binding sites contribute to the evidenced activation. What is known is that compared to the full forward NSE, the full reverse NSE is significantly activated by the combined action of Sox2 and Wnt pathway effectors Lef1-βcatenin.

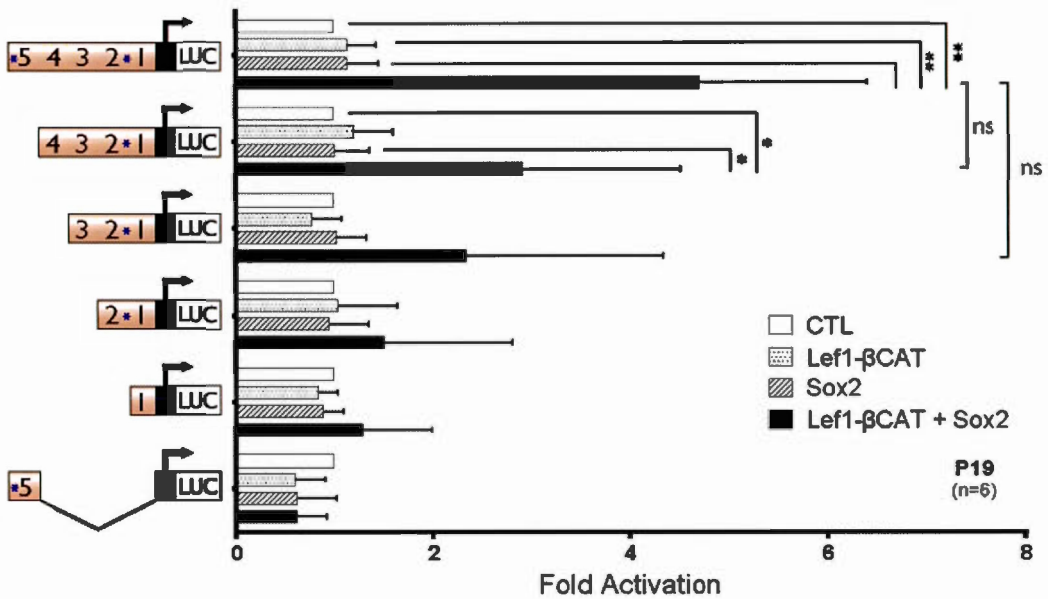


Figure 5.4 Fragmentation results of *Cdx2*NSEreverse.
Dose was 100ng of activator DNA and 100ng of *Cdx2*NSEreverse fragments. Lef1-βcatenin and Sox2 together robustly activated the promoter, of the complete enhancer NSE reporter followed by the second most complete enhancer (BS4,3,2,1). In all there was a trend present, where the bigger the enhancer the bigger the activation. An asterisk (**) denotes a very significant $p \leq 0.005$; and (*) a significant $p \leq 0.05$. The Blue asterisk (*) marks the presence of the composed STEM binding site that includes Sox2.

CHAPTER VI

DISCUSSION

6.1 In pursuit of a conditional *Cdx1* overexpression mouse model

Homologous recombination has long been used to insert exogenous DNA constructs into a specific genomic locus in mice (Capecchi, 2005; Misra and Duncan, 2002). The *ROSA26* locus in particular has been established as the preferred genomic locus for gene targeting (Casola, 2010; Nyabi et al., 2009), because transgene integration is stable, and expression levels are constant and ubiquitous, not to mention that disturbing the locus produces no phenotype (Soriano, 1999; Zambrowicz et al., 1997). New targeting strategies have emerged that successfully target the *ROSA26* locus in mice. These include zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered regularly interspaced short palindromic repeats CRISPR/Cas RNA-guided nucleases (Cas9/gRNA) (Hermann et al., 2012; Hsu et al., 2014; Kasparek et al., 2014; Meyer et al., 2010; Sato et al., 2015). These nuclease based technologies are capable of cleaving genomic DNA at specific sites, causing double-stranded breaks to improve homologous recombination as well as insertion of exogenous constructs mediated by DNA repair. They are timesaving and cost-efficient methods that have as high as 4.5% (ZFNs), 8.8% (TALENs), and 35% (Cas9/gRNA) recombination efficiency in gene targeting of the mouse *ROSA26* locus (Fujii et al., 2013; Kasparek et al., 2014; Meyer et al., 2010). However, there is still some uncertainty regarding possible off-target DNA cleavages caused by these new nuclease-based strategies, as well as the cytotoxicity that could be brought about by the double-stranded breaks. Targeting based on TALENs in particular seem most

promising, with Kasparek et al. (2014) reporting no off-target activity found in their approach. Nevertheless, concerns remain since off-target DNA cleavages can still be induced by these methods, resulting in unknown mutations that are difficult to detect (Kong et al., 2014; Koo et al., 2015; Owens et al., 2013). In the end, the traditional homologous recombination strategy aimed at targeting the *ROSA26* locus is still the established solution, and our preferred choice, since it does not induce off-target mutations.

We targeted the mouse *ROSA26* locus with a vector containing a floxed-PGK-neo-tpA stop cassette upstream of *FLAG-Cdx1* and *IRES-eGFP*. After transfection and G418 selection, selected colonies were screened for the knock-in allele. Out of 96 clones screened, 3 (3.1%) were PCR-positive for integration of the knock-in construct corresponding to the F1-R1 primer screen parameters. Southern blot was inconclusive, with none of the controls working, indicating that technical issues most likely played a role. It would be very interesting to find out how often PCR-positive *ROSA26* locus targeted clones are also Southern blot-positive. Soriano (1999) reported that 8 out of 8 (100%) were. Of note, our 1.1-kb PCR screen method was nearly identical in location as described by (1.2-kb) Soriano (1999). Despite the more-than 103 *ROSA26* locus knock-ins described to date (Casola, 2010), this information is now likely seen as trivial and no longer divulged. Moreover, if any verification details are given they usually extend to Southern blot confirmation. Southern blot not only validates 5' and 3' transgene insertion but also distinguishes the wild-type from the targeted allele. The advantage of PCR is that it allows us to examine a large number of clones in a high-throughput manner, sifting and identifying a smaller number of possibly positive clones. In the end, even if our three clones were validated by southern blot, that is just 3.1% of all clones, which is below the reported average targeting efficiency for the *ROSA26* locus, around 25% for G418 selected clones (Hohenstein et al., 2008). The lower number of positive clones could be to due to transfection optimization issues.

In this project, 96 drug-selected clones from a 96-well feeder plate were part transferred onto four 24-well tissue culture plates, and the clones remaining in the 96-well plate were then stored at -80°C. Cells in the 24-well plates were allowed to expand until confluent. The advantage of transferring 96 drug-selected clones onto four 24-well plates versus another 96-well plate is that the cell lysate from a 24-well plate permits for several Southern blots whereas a 96-well plate would only produce enough DNA for one Southern blot (Limaye et al., 2009). In our case, there was enough DNA for two southern blots targeting the 5' and 3' *ROSA26* knock-in locus. However, since there was a need for more DNA, the entire 96-well plate used for storing the clones at -80°C was defrosted and the three PCR-positive clones and one WT clone was passaged for subsequent storing and DNA harvesting. Once again, the harvested DNA was exhausted in two more Southern blots that failed to target the 5' and 3' knock-in locus, and WT *ROSA26* locus. For future work this suggests that it may be best to start back from the targeting vector, instead of defrosting and passaging more clones. For one, passaging ES cells more and more increases the potential for differentiation. Two, storing the ES clones—now cryogenically frozen in liquid nitrogen—for long periods of time diminishes their capacity to be successfully microinjected. And three, literature indicates that targeting the *ROSA26* locus is highly efficient. This means that improvements in transfection conditions could likely yield more positive clones, closer to the reported 25% average (Hohenstein et al., 2008). In addition, no control probes were tested. It may be worthwhile to target the GFP sequence, for example.

If homologous recombination and integration of our *Cdx1* transgene were confirmed, we would be well on our way to generate a knock-in mouse capable of conditional *Cdx1* overexpression. Knowing that Cdx proteins play a significant role in the initial steps of NC development and that overexpression of genes needed for NC development is manifested in NC derived tumours (Trainor, 2013), we had hypothesized that a mouse line overexpressing *Cdx1* would generate NCC derived

tumours such as neuroblastoma. Without the mouse line, we are still left with this question. Moreover, still to be discovered are the regulatory mechanisms implicating the neural functions of Cdx1 proteins to the supposed NC-derived malignancy. Of note, this mouse line capable of conditional *Cdx1* overexpression would be the first to demonstrate such a link, with overexpression particularly aimed at neural crest development. There exists one reported transgenic *Cdx1* overexpression mouse model whose expression is not inducible in specific tissues/context and is under the control of a *Cdx1* promoter instead of the stable and ubiquitous expression conferred by the *ROSA26* promoter; overexpression in this case had homeotic and axial patterning defects (Gaunt et al., 2008).

6.2 Insights into *Cdx2*NSE neural regulation and potential application

Cdx1 genes are just one of three *Cdx* genes that have been maintained by vertebrates for over hundreds of millions of years, namely *Cdx1/2/4*[*] in mammals, birds, amphibians and certain fish (Faas and Isaacs, 2009; Marletaz et al., 2015; Mulley and Holland, 2010). Cdx proteins share some spatial overlap as well as core pathways, particularly those necessary for AP vertebral patterning (Beck and Stringer, 2010). There is also evidence indicating that Cdx proteins mediate *Pax3* expression—Pax3 being a key transcription factor in the NC-GRN (Sanchez-Ferras et al., 2014; Sanchez-Ferras et al., 2012). See Appendix B for our study on *Pax3* regulation by Cdx2 proteins. Prior to the emergence of the NC-GRN, pre-vertebrate chordates likely only had one *Cdx* gene, as is the case for the cephalochordate amphioxus (*AmphiCdx* gene) (Brooke et al., 1998). Cdx1/2/4 proteins may share similar targets, and some redundancy may be an evolutionary safeguard advantage, but Marletaz et al. (2015) argues that while corroboration and overlap exist between Cdx proteins,

[*] Orthologous gene nomenclature may differ between species

their individual functional roles are likely distinct. This may indicate that prior to the emergence of Cdx1/2/4, the sole *Cdx* gene precursor underwent gene duplication and sequential modifications that lead to three neo-functionalized Cdx protein-coding genes. It is no secret that sequential genetic modifications guide the evolution of novel features (Holland et al., 1994; Pennisi, 2008). However, whether the critical driving force underlying vertebrate evolution was due to protein neo-functionalization or changes in cis-regulatory sequences is still contested (Carroll, 2008; Levine, 2010; Lynch and Wagner, 2008). Concerning the evolution of vertebrates and the NC-GRN, both Van Otterloo et al. (2013) and Jandzik et al. (2015) make the case that mutations in pre-vertebrate cis-regulatory regions were critical for NC evolution. Van Otterloo et al. (2013) go so far as to suggest that there is little evidence supporting the claim that protein neo-functionalization guided the evolution of the NC-GRN. This line of thinking is the most accepted by evolutionary developmental biologists (Hoekstra and Coyne, 2007; Wray, 2007). Part of the reasoning is that many protein-coding genes have highly conserved functions among different species (such as non-vertebrate *Cdx* gene homologs *AmphiCdx* and *cad*) but variations in expression levels abound (Wittkopp and Kalay, 2012). Whether the driving force for NC-GRN evolution was due to protein-coding or cis-regulatory changes, both evolutionary novelties have ultimately contributed to its development.

In the case of *Cdx*2NSE, the enhancer is undoubtedly a vertebrate-exclusive feature, involving a cis-regulatory sequence that enhances the expression of a vertebrate "neo-functionalized" protein. Clearly it is difficult to discern, without going back in time, which novel feature was more crucial for vertebrate development, the (NSE) enhancer for *Cdx2* or *Cdx2* itself. Of course the main contention is not whether one feature came first or not, but which novelty was critical for vertebrate evolution (and possibly for evolution of Cdx neural functions). If we follow convention, most evolutionary biologists would probably agree that divergences in the NSE sequence are more likely to guide important phenotypic changes. However, the role and impact

of the NSE is not yet known. A vertebrate homology comparison of *Cdx2*NSE carried out by Wang and Shashikant (2007) indicated a 65–98% NSE sequence similarity between mouse and *Xenopus*, chick, opossum, dog, rat and human. There was no homology with zebrafish and fugu, which is not surprising given the fact that teleost fish, like zebrafish and fugu, do not have *Cdx2* genes, but instead have two copies of *Cdx1* and a single *Cdx4* gene (Mulley et al., 2006). Both *Cdx2* and its own NSE seem to be highly conserved together, which indicates a selective pressure to maintain both features.

While our take in the studying the enhancer does not come from an evolutionary standpoint but a practical one, it is worth noting that enhancers as a whole are often highly conserved, yet subtle variations are likely between species and thus we cannot rely on homology alone when analyzing putative TFBSs. Our own homology comparison of the 852-bp NSE, described by Wang and Shashikant (2007), with 100-vertebrate species from the UCSC Genome browser also validated the high conservation of the NSE. However, there was divergence in sequence conservation for putative Lef1/Tcf bindings sites found by computational analysis, with only putative sites BS2, BS3, and BS5 appearing to have conservation within the 100-vertebrates. A more selective search, including just mouse, human, rhesus and elephant does reveal conservation of putative BS4 but not BS1, which is the most divergent of all five potential Lef1/Tcf binding sites. Luciferase assay tests of BS1 by itself, with canonical Wnt pathway effectors Lef1-βcatenin revealed no activation, even when co-transfected with Sox2. Granted, the synergy of both Lef1-βcatenin and Sox2 was only evident in the full *Cdx2*NSEreverse-luc reporter, thus no real inferences can be made on individual or groups of potential binding sites. What can be said is that a trend is present between the enhancer fragments (Figure 5.4), where the larger the fragment the greater the activation. Also, it is unclear if the two highly conserved potential STEM (Sox2) binding sites contributed to any activation.

The activity of *Cdx2*NSE is not entirely located in the neural specific regions of the developing posterior neuroectoderm and NT, but also slightly in posterior mesodermal tissue (Coutaud and Pilon, 2013). The goal here was to test the NSE for neural specific activity, and identify the regions most receptive to neural specific activation so that a truly neural specific enhancer could be constructed and used for the generation of a Cre mouse line driven by the new enhancer. This new modified enhancer would hopefully be both neural specific and innocuous *in vivo*. Even if luciferase assays had identified a promising NSE, its true utility would not be apparent until tested in the transgenic mouse line. This risk may be worth taking considering how such a tool, aimed at posterior neural development, does not yet exist.

Since no one piece of the NSE architecture stood out more in neural specific activation, it remains to be identified. If, for example, one small segment had been able to nearly recapitulate the neural specific activation of the full sequence, then by exclusion the missing segments could be categorized as non-neural specific. While still just a clue this could indicate that the smaller segment strongly sensitive to neural specific activation could be used in regenerating a truly *Cdx*2NSE Cre mouse line. Unable to distinguish these sites through activation by Sox2 and Lef1-βcatenin, there are two suggestions for future research. One is to repeat, and two is to test NSE activation by other activator proteins needed for posterior neurogenesis.

Admittedly, concerning the discrepancy between the activation displayed by the full *Cdx2*NSEreverse in tests from Figure 5.4 and 5.3, it would have been better to maintain the same cell density as before, for continuity, and reproducibility. As mentioned in materials and methods, the first tests demonstrated in Figure 5.3 used $3x10^4$ cells/well, and $1.5x10^4$ cells/well in the second tests shown in Figure 5.4. Somewhere along the way $3x10^4$ cells/well mistakenly became $3x10^4$ cells/mL and thus $1.5x10^4$ cells/well (i.e. 500µl/well). Still, $3x10^4$ cells/mL falls somewhat close to

the recommended (Novagen) density range of 4-8x10^4 cells/mL for adherent cells—ranging from three-quarters to about one-third the amount. Even if not the recommended amount, one (n=1) 6x10^4 cells/mL was compared to one (n=1) 3x10^4 cells/mL, using the same transfection materials and conditions (done on the same day). The results did not suggest more cells were responsible for the discrepancy, in fact slightly more activation was observed in the test with less cell density. This could suggest the inconsistency is due in large part to lower quality of transfection materials, as the substantial time lapse from the first test and the second test meant the original materials had been exhausted and new ones were made. The comparison being an n=1 is of course not significant and a rework is still warranted, yet it is clear that materials and conditions need to be as similar as possible, and the best strategy is to test consecutively in as short a period of time with enough of the same resources for the whole duration. Despite the oversight, conditions were kept the same for the second part of the tests, and so all things being equal, inferences can still be made, though again still warrants retesting.

Testing the NSE with other possible key activators, and their combinations, could also help identify regions more sensitive to neural specific activation. MatInspector analysis identified four Ets motifs involved in FGF signalling (Appendix C). The most conserved of the four motifs were the two most-inner ones. Like with the putative Lef1/Tcf motifs involved in Wnt signalling, the Ets motifs were mostly detected on the reverse NSE strand (3 out of 4). It could be the case that activation by both Wnt and FGF pathways is synergistic.

Cdx2 has also been shown to regulate its own expression, through response elements, sensitive to Cdx2 activation, located proximal to the promoter (Xu et al., 1999). However, MatInspector did not identify any motifs for Cdx2 (or Cdx1/4).

While the Sox-B1 transcription family of transcription factors Sox1, Sox2 and Sox3 share 80% sequence similarity and are functionally redundant (Zhang, 2014), Sox2

remains the most definite pan-neural specific marker of the early neural plate (Okuda et al., 2010). In addition, Sox2 plays a role in maintaining neural progenitor populations throughout early development and beyond (Ellis et al., 2004). Sox2 along with Oct4 and Nanog play a role in maintaining cell pluripotency as well as specifying lineages. Sox2 maintains neuroectodermal identity while Nanog and Oct4 direct the differentiation of the mesoendoderm (Loh and Lim, 2011; Thomson et al., 2011; Zhang, 2014). Sox2, Nanog and Oct4 are all part of the two putative composed STEM binding sites found on the reverse *Cdx2*NSE strand. It could be possible that these sites are equally responsible for the neural specific and mesoderm activity of the enhancer observed by Coutaud and Pilon (2013). If true, separating two fate-guiding traits from one shared binding site is likely to be unmanageable.

Enhancers are robust and complex regulatory features. Unlike mutations in protein-coding sequences, changes to enhancers are not as deleterious, nor likely to be more pleiotropic (Wittkopp and Kalay, 2012). Yet, their contribution to evolution is believed to be major. Attempting to uncover and refashion the complexity and robustness acquired through hundreds of millions of years of evolution is a formidable task. The identification of neural specific regions rests to be discovered, and the hope is that, despite the challenges, retesting and adding the action of FGF signalling will help uncover these regions.

APPENDIX A

EVOLUTIONARY CONSERVATION OF *CDX2*NSE POTENTIAL TFBS

Table A.1 Potential TFBS phyloP plot scores.
Core sequences are the most conserved four consecutive nucleotides. Positive scores indicate conservation, whereas negative scores predict acceleration. Data retrieved from the UCSC genome browser.

Query Sequence	Bases	Minimum	Maximum	Mean	Standard Deviation
Cdx2NSE	842	-5.114	4.8243	1.02275	1.71358
BS1	17	-1.70244	0.439701	-0.381687	0.608727
BS1 CORE	4	-0.433024	0.122346	-0.195008	0.282368
STEM1	19	0.122346	4.16861	1.90538	1.31179
STEM1 CORE	4	0.836394	2.42317	1.82813	0.690139
BS2	17	1.94713	4.80332	3.65525	0.9985
BS2 CORE	4	3.29589	4.80332	3.83143	0.71368
BS3	16	-0.414598	3.91757	1.20367	1.17115
BS3 CORE	4	-0.112354	1.90261	0.970687	0.852398
BS4	17	-2.12732	2.3056	0.261006	1.03406
BS4 CORE	4	-0.515346	0.0891417	-0.112354	0.284958
BS5	21	-0.716843	3.01083	1.49482	1.09648
BS5 CORE1	4	1.39887	3.01083	2.40635	0.698003
BS5 CORE2	4	1.29812	3.01083	1.85223	0.791155
STEM2	19	-0.213102	3.11158	1.66399	0.976333
STEM2 CORE	4	0.995874	3.11158	1.70111	0.95578

APPENDIX B

REGULATION OF *PAX3* NCE3 BY CDX2 PROTEINS

The paired box transcription factor Pax3 is an important regulator of the NC-GRN circuit, involved in the specification of the neural crest. There are three neural crest element (NCE) enhancers for *Pax3*: NCE1, NCE2 and NCE3. Recent enhancer studies have demonstrated that Wnt-mediated signalling activates NCE2 through intermediary Cdx proteins, and that Zic2 transcription factors also regulate NCE2 (Sanchez-Ferras et al., 2014; Sanchez-Ferras et al., 2012). NCE1/2 were the first neural crest enhancers to be described, and as most enhancers go, they are highly conserved; they are also similar in function (Li et al., 1999; Milewski et al., 2004; Pruitt et al., 2004). The third enhancer known in the lab, as NCE3, was discovered by Degenhardt et al. (2010) and is an intronic enhancer with seemingly redundant functions to NCE1/2. Therefore, the hypothesis is that Cdx proteins regulate the NCE3 of *Pax3* just as other enhancers of *Pax3*. Presented here are the luciferase tests analyzing the effect of Cdx2 as an activator on NCE3 forward and reverse sequences that we inserted into a luciferase expression vector with an 800-bp *Pax3* promoter. There seems to be dose-dependent activation of the *Pax3* promoter by Cdx2 proteins with maximum activation achieved at Cdx2 dose 25ng (Figure A.1).

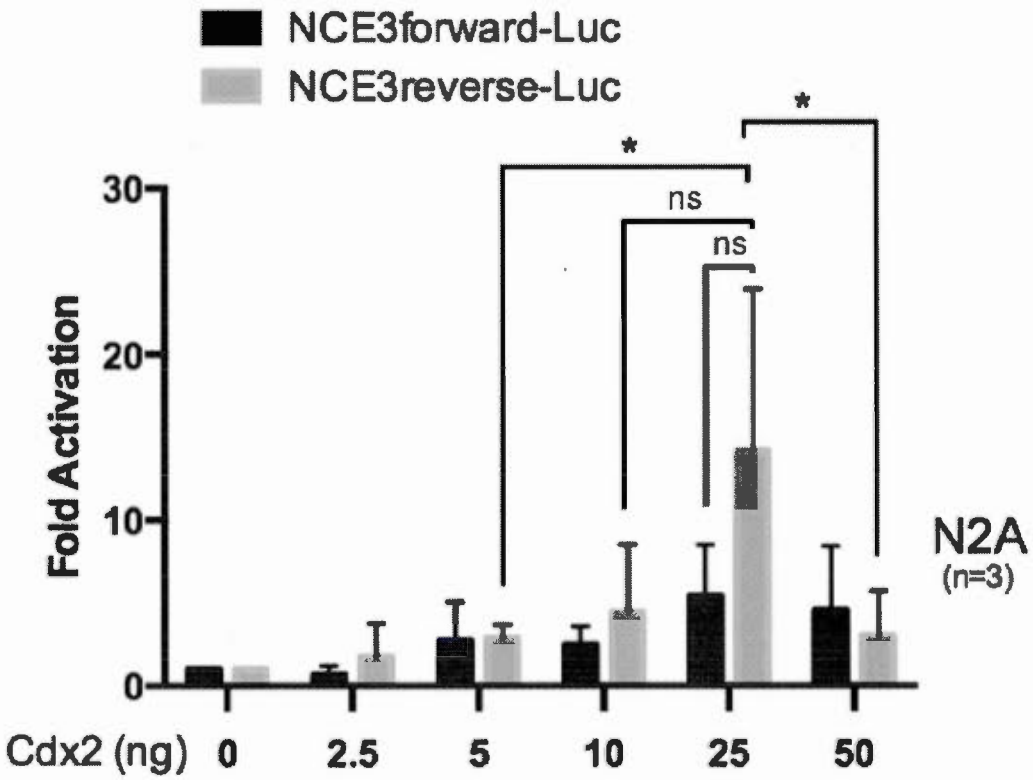


Figure B.1 Effects on NCE3 reporter activation by Cdx2 proteins.
(*) represents $p \leq 0.05$

APPENDIX C

*CDX2*NSE AND FGF SIGNALLING

Check transcription factor <-> matrix family assignment				Position			Strand	Matrix sim.	Additional lines of evidence		Sequence (red: ci-value > 60 CAPITALS: core sequence)
Matrix Family	Detailed Family Information	Matrix	Detailed Matrix Information	from	to	anchor			#	Evidence	
V$ETSF	Human and murine ETS1 factors	V$SPI1.03	SPI-1 proto-oncogene; hematopoietic transcription factor PU.1	217	237	227	(-)	0.901	0		cctgctttGGAAgtaaagggga
V$ETSF	Human and murine ETS1 factors	V$GABP.01	GABP: GA binding protein	301	321	311	(+)	0.893	0		cctggcggGGAAgggcgtgaa
V$ETSF	Human and murine ETS1 factors	V$SPIB.01	Spi-B transcription factor (Spi-1/PU.1 related)	430	450	440	(-)	0.924	0		ctgaaaggGGAAcacagtcgc
V$ETSF	Human and murine ETS1 factors	V$ELF2.01	Ets - family member ELF-2 (NERF1a)	655	675	665	(-)	0.900	0		gctggccaGGAActgttccac

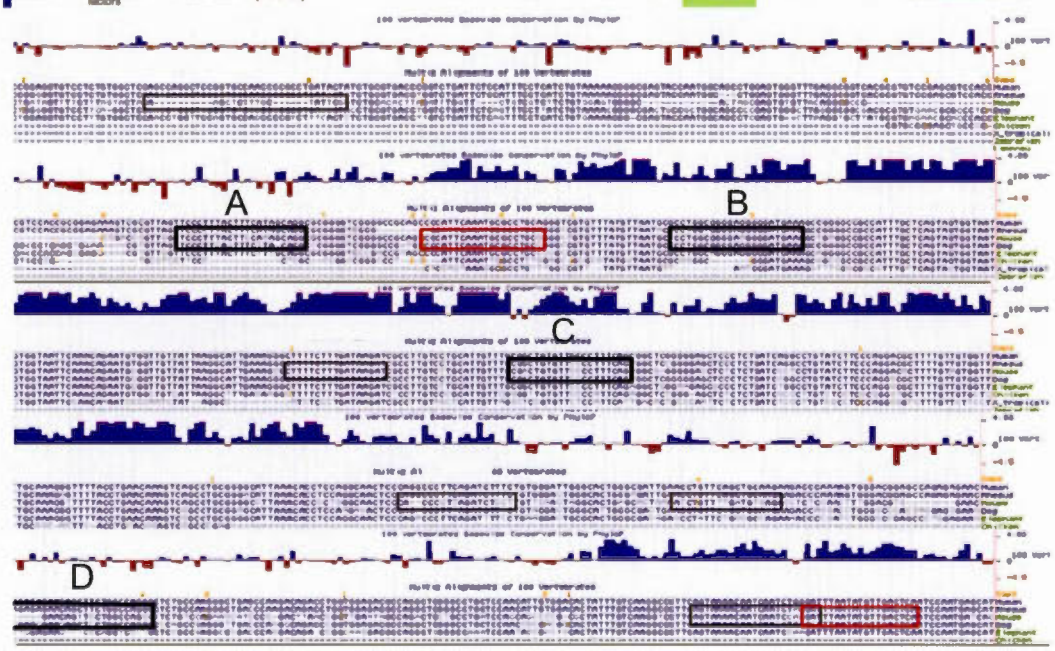


Figure C.2 Placement of putative Ets motifs and homology.
Ets motifs involve FGF signalling. GGAA characterize the core sequences of most Ets motifs, which makes their importance more problematic to address. At the top are the results of MatInspector; there are four putative Ets motifs, with high matrix similarity particularly for the three reverse NSE stands. At the bottom are the results of 100-vertebrate homology comparison (UCSC genome browser) with respect to putative binding sites. A-C are the putative Ets motifs, in red are the putative motifs for the STEM binding sites, and light black the putative Lef1/Tcf binding sites. Ets sites B/C are the most conserved.

REFERENCES

Akiyama, H., Chaboissier, M.C., Behringer, R.R., Rowitch, D.H., Schedl, A., Epstein, J.A., de Crombrugghe, B., 2004. Essential role of Sox9 in the pathway that controls formation of cardiac valves and septa. Proc Natl Acad Sci U S A 101, 6502-6507.

Althoff, K., Beckers, A., Bell, E., Nortmeyer, M., Thor, T., Sprussel, A., Lindner, S., De Preter, K., Florin, A., Heukamp, L.C., Klein-Hitpass, L., Astrahantseff, K., Kumps, C., Speleman, F., Eggert, A., Westermann, F., Schramm, A., Schulte, J.H., 2015. A Cre-conditional MYCN-driven neuroblastoma mouse model as an improved tool for preclinical studies. Oncogene 34, 3357-3368.

Amore, G., Yavrouian, R.G., Peterson, K.J., Ransick, A., McClay, D.R., Davidson, E.H., 2003. Spdeadringer, a sea urchin embryo gene required separately in skeletogenic and oral ectoderm gene regulatory networks. Dev Biol 261, 55-81.

Angers, S., Moon, R.T., 2009. Proximal events in Wnt signal transduction. Nat Rev Mol Cell Biol 10, 468-477.

Antony, A.C., Hansen, D.K., 2000. Hypothesis: folate-responsive neural tube defects and neurocristopathies. Teratology 62, 42-50.

Auerbach, R., 1954. Analysis of the developmental effects of a lethal mutation in the house mouse. Journal of Experimental Zoology 127, 305-329.

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

Baird, P.A., Sadovnick, A.D., 1984. Survival in infants with anencephaly. Clin Pediatr (Phila) 23, 268-271.

Barad, M., Jack, T., Chadwick, R., McGinnis, W., 1988. A novel, tissue-specific, Drosophila homeobox gene. EMBO J 7, 2151-2161.

Beck, F., Erler, T., Russell, A., James, R., 1995. Expression of Cdx-2 in the mouse embryo and placenta: possible role in patterning of the extra-embryonic membranes. Dev Dyn 204, 219-227.

Beck, F., Stringer, E.J., 2010. The role of Cdx genes in the gut and in axial development. Biochem Soc Trans 38, 353-357.

Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W.J., Mattick, J.S., Haussler, D., 2004. Ultraconserved elements in the human genome. Science 304, 1321-1325.

Bel-Vialar, S., Itasaki, N., Krumlauf, R., 2002. Initiating Hox gene expression: in the early chick neural tube differential sensitivity to FGF and RA signaling subdivides the HoxB genes in two distinct groups. Development 129, 5103-5115.

Beland, M., Pilon, N., Houle, M., Oh, K., Sylvestre, J.R., Prinos, P., Lohnes, D., 2004. Cdx1 autoregulation is governed by a novel Cdx1-LEF1 transcription complex. Mol Cell Biol 24, 5028-5038.

Brazel, C.Y., Limke, T.L., Osborne, J.K., Miura, T., Cai, J., Pevny, L., Rao, M.S., 2005. Sox2 expression defines a heterogeneous population of neurosphere-forming cells in the adult murine brain. Aging Cell 4, 197-207.

Britsch, S., Goerich, D.E., Riethmacher, D., Peirano, R.I., Rossner, M., Nave, K.-A., Birchmeier, C., Wegner, M., 2001. The transcription factor Sox10 is a key regulator of peripheral glial development. Genes & development 15, 66-78.

Brodeur, G.M., 2003. Neuroblastoma: biological insights into a clinical enigma. Nat Rev Cancer 3, 203-216.

Brooke, N.M., Garcia-Fernandez, J., Holland, P.W., 1998. The ParaHox gene cluster is an evolutionary sister of the Hox gene cluster. Nature 392, 920-922.

Capecchi, M.R., 2005. Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. Nat Rev Genet 6, 507-512.

Carroll, S.B., 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell 134, 25-36.

Casola, S., 2010. Mouse models for miRNA expression: the ROSA26 locus. Methods Mol Biol 667, 145-163.

Chawengsaksophak, K., de Graaff, W., Rossant, J., Deschamps, J., Beck, F., 2004. Cdx2 is essential for axial elongation in mouse development. Proc Natl Acad Sci U S A 101, 7641-7645.

Colas, J.F., Schoenwolf, G.C., 2001. Towards a cellular and molecular understanding of neurulation. Dev Dyn 221, 117-145.

Conway, S.J., Henderson, D.J., Kirby, M.L., Anderson, R.H., Copp, A.J., 1997. Development of a lethal congenital heart defect in the splotch (Pax3) mutant mouse. Cardiovascular research 36, 163-173.

Copp, A.J., Greene, N.D., 2013. Neural tube defects--disorders of neurulation and related embryonic processes. Wiley Interdiscip Rev Dev Biol 2, 213-227.

Copp, A.J., Greene, N.D., Murdoch, J.N., 2003. The genetic basis of mammalian neurulation. Nat Rev Genet 4, 784-793.

Coutaud, B., 2013. Génération d'un nouvel outil pour l'étude in vivo de la neurogenèse et caractérisation de la régulation et des fonctions neurales des protéines CDX (Mémoire de maîtrise). Université du Québec à Montréal. Récupéré d'*Archipel*, l'archive de publications électroniques de l'UQAM http://www.archipel.uqam.ca/5914/.

Coutaud, B., Pilon, N., 2013. Characterization of a novel transgenic mouse line expressing Cre recombinase under the control of the Cdx2 neural specific enhancer. Genesis 51, 777-784.

Czeizel, A.E., Bartfai, Z., Banhidy, F., 2011. Primary prevention of neural-tube defects and some other congenital abnormalities by folic acid and multivitamins: history, missed opportunity and tasks. Ther Adv Drug Saf 2, 173-188.

Danielian, P.S., Muccino, D., Rowitch, D.H., Michael, S.K., McMahon, A.P., 1998. Modification of gene activity in mouse embryos in utero by a tamoxifen-inducible form of Cre recombinase. Curr Biol 8, 1323-1326.

Degenhardt, K.R., Milewski, R.C., Padmanabhan, A., Miller, M., Singh, M.K., Lang, D., Engleka, K.A., Wu, M., Li, J., Zhou, D., Antonucci, N., Li, L., Epstein, J.A., 2010. Distinct enhancers at the Pax3 locus can function redundantly to regulate neural tube and neural crest expressions. Dev Biol 339, 519-527.

Dixon, J., Jones, N.C., Sandell, L.L., Jayasinghe, S.M., Crane, J., Rey, J.P., Dixon, M.J., Trainor, P.A., 2006. Tcof1/Treacle is required for neural crest cell formation and proliferation deficiencies that cause craniofacial abnormalities. Proc Natl Acad Sci U S A 103, 13403-13408.

Draper, B.W., Stock, D.W., Kimmel, C.B., 2003. Zebrafish fgf24 functions with fgf8 to promote posterior mesodermal development. Development 130, 4639-4654.

Drewett, V., Molina, H., Millar, A., Muller, S., von Hesler, F., Shaw, P.E., 2001. DNA-bound transcription factor complexes analysed by mass-spectrometry: binding of novel proteins to the human c-fos SRE and related sequences. Nucleic Acids Res 29, 479-487.

Duband, J.L., 2010. Diversity in the molecular and cellular strategies of epithelium-to-mesenchyme transitions: Insights from the neural crest. Cell Adh Migr 4, 458-482.

Dubois, N.C., Hofmann, D., Kaloulis, K., Bishop, J.M., Trumpp, A., 2006. Nestin-Cre transgenic mouse line Nes-Cre1 mediates highly efficient Cre/loxP mediated recombination in the nervous system, kidney, and somite-derived tissues. Genesis 44, 355-360.

Dudas, M., Sridurongrit, S., Nagy, A., Okazaki, K., Kaartinen, V., 2004. Craniofacial defects in mice lacking BMP type I receptor Alk2 in neural crest cells. Mech Dev 121, 173-182.

Dupin, E., Calloni, G.W., Le Douarin, N.M., 2010. The cephalic neural crest of amniote vertebrates is composed of a large majority of precursors endowed with neural, melanocytic, chondrogenic and osteogenic potentialities. Cell Cycle 9, 238-249.

Dupin, E., Sommer, L., 2012. Neural crest progenitors and stem cells: from early development to adulthood. Dev Biol 366, 83-95.

Durston, A.J., 2015. Time, space and the vertebrate body axis. Semin Cell Dev Biol.

Echelard, Y., Vassileva, G., McMahon, A.P., 1994. Cis-acting regulatory sequences governing Wnt-1 expression in the developing mouse CNS. Development 120, 2213-2224.

Ellis, P., Fagan, B.M., Magness, S.T., Hutton, S., Taranova, O., Hayashi, S., McMahon, A., Rao, M., Pevny, L., 2004. SOX2, a persistent marker for multipotential neural stem cells derived from embryonic stem cells, the embryo or the adult. Dev Neurosci 26, 148-165.

Endo, Y., Osumi, N., Wakamatsu, Y., 2002. Bimodal functions of Notch-mediated signaling are involved in neural crest formation during avian ectoderm development. Development 129, 863-873.

Epstein, D.J., 2009. Cis-regulatory mutations in human disease. Brief Funct Genomic Proteomic 8, 310-316.

Epstein, D.J., Malo, D., Vekemans, M., Gros, P., 1991a. Molecular characterization of a deletion encompassing the splotch mutation on mouse chromosome 1. Genomics 10, 89-93.

Epstein, D.J., Vekemans, M., Gros, P., 1991b. Splotch (Sp2H), a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of Pax-3. Cell 67, 767-774.

Erickson, J.D., 2002. Folic acid and prevention of spina bifida and anencephaly. 10 years after the U.S. Public Health Service recommendation. MMWR Recomm Rep 51, 1-3.

Faas, L., Isaacs, H.V., 2009. Overlapping functions of Cdx1, Cdx2, and Cdx4 in the development of the amphibian Xenopus tropicalis. Dev Dyn 238, 835-852.

Feuchtbaum, L.B., Currier, R.J., Riggle, S., Roberson, M., Lorey, F.W., Cunningham, G.C., 1999. Neural tube defect prevalence in California (1990-1994): eliciting patterns by type of defect and maternal race/ethnicity. Genet Test 3, 265-272.

Fletcher, J.M., Brei, T.J., 2010. Introduction: Spina bifida--a multidisciplinary perspective. Dev Disabil Res Rev 16, 1-5.

Fujii, W., Kawasaki, K., Sugiura, K., Naito, K., 2013. Efficient generation of large-scale genome-modified mice using gRNA and CAS9 endonuclease. Nucleic Acids Res 41, e187.

Gamer, L.W., Wright, C.V., 1993. Murine Cdx-4 bears striking similarities to the Drosophila caudal gene in its homeodomain sequence and early expression pattern. Mech Dev 43, 71-81.

Garnett, A.T., Square, T.A., Medeiros, D.M., 2012. BMP, Wnt and FGF signals are integrated through evolutionarily conserved enhancers to achieve robust expression of Pax3 and Zic genes at the zebrafish neural plate border. Development 139, 4220-4231.

Gaunt, S.J., Drage, D., Trubshaw, R.C., 2008. Increased Cdx protein dose effects upon axial patterning in transgenic lines of mice. Development 135, 2511-2520.

Ghysen, A., 2003. The origin and evolution of the nervous system. Int J Dev Biol 47, 555-562.

Gomez, C., Ozbudak, E.M., Wunderlich, J., Baumann, D., Lewis, J., Pourquie, O., 2008. Control of segment number in vertebrate embryos. Nature 454, 335-339.

Goulding, M., Sterrer, S., Fleming, J., Balling, R., Nadeau, J., Moore, K., Brown, S., Steel, K., Gruss, P., 1993. Analysis of the Pax-3 gene in the mouse mutant splotch. Genomics 17, 355-363.

Goulding, M.D., Chalepakis, G., Deutsch, U., Erselius, J.R., Gruss, P., 1991. Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. EMBO J 10, 1135-1147.

Graham, V., Khudyakov, J., Ellis, P., Pevny, L., 2003. SOX2 functions to maintain neural progenitor identity. Neuron 39, 749-765.

Guthrie, S., Lumsden, A., 1991. Formation and regeneration of rhombomere boundaries in the developing chick hindbrain. Development 112, 221-229.

Hall, B.K., 1999. The neural crest in development and evolution. Springer Science & Business Media.

Harris, M.L., Buac, K., Shakhova, O., Hakami, R.M., Wegner, M., Sommer, L., Pavan, W.J., 2013. A dual role for SOX10 in the maintenance of the postnatal melanocyte lineage and the differentiation of melanocyte stem cell progenitors. PLoS Genet 9, e1003644.

Hendrickx, M., Van, X.H., Leyns, L., 2009. Anterior-posterior patterning of neural differentiated embryonic stem cells by canonical Wnts, Fgfs, Bmp4 and their respective antagonists. Dev Growth Differ 51, 687-698.

Hermann, M., Maeder, M.L., Rector, K., Ruiz, J., Becher, B., Burki, K., Khayter, C., Aguzzi, A., Joung, J.K., Buch, T., Pelczar, P., 2012. Evaluation of OPEN zinc finger nucleases for direct gene targeting of the ROSA26 locus in mouse embryos. PLoS One 7, e41796.

Hikasa, H., Sokol, S.Y., 2013. Wnt signaling in vertebrate axis specification. Cold Spring Harb Perspect Biol 5, a007955.

Hinman, V.F., Nguyen, A.T., Cameron, R.A., Davidson, E.H., 2003. Developmental gene regulatory network architecture across 500 million years of echinoderm evolution. Proc Natl Acad Sci U S A 100, 13356-13361.

Hoekstra, H.E., Coyne, J.A., 2007. The locus of evolution: evo devo and the genetics of adaptation. Evolution; international journal of organic evolution 61, 995-1016.

Hohenstein, P., Slight, J., Ozdemir, D.D., Burn, S.F., Berry, R., Hastie, N.D., 2008. High-efficiency Rosa26 knock-in vector construction for Cre-regulated overexpression and RNAi. PathoGenetics 1, 3.

Holland, P.W., Garcia-Fernàndez, J., Williams, N.A., Sidow, A., 1994. Gene duplications and the origins of vertebrate development. Development 1994, 125-133.

Hong, C.S., Park, B.Y., Saint-Jeannet, J.P., 2008. Fgf8a induces neural crest indirectly through the activation of Wnt8 in the paraxial mesoderm. Development 135, 3903-3910.

Houart, C., Caneparo, L., Heisenberg, C., Barth, K., Take-Uchi, M., Wilson, S., 2002. Establishment of the telencephalon during gastrulation by local antagonism of Wnt signaling. Neuron 35, 255-265.

Houle, M., Allan, D., Lohnes, D., 2003a. Cdx homeodomain proteins in vertebral patterning, Advances in Developmental Biology and Biochemistry. Elsevier, pp. 69-105.

Houle, M., Sylvestre, J.R., Lohnes, D., 2003b. Retinoic acid regulates a subset of Cdx1 function in vivo. Development 130, 6555-6567.

Howard, M.L., Davidson, E.H., 2004. cis-Regulatory control circuits in development. Dev Biol 271, 109-118.

Hsu, P.D., Lander, E.S., Zhang, F., 2014. Development and applications of CRISPR-Cas9 for genome engineering. Cell 157, 1262-1278.

Hu, N., Strobl-Mazzulla, P.H., Bronner, M.E., 2014. Epigenetic regulation in neural crest development. Dev Biol 396, 159-168.

Huang, X., Saint-Jeannet, J.P., 2004. Induction of the neural crest and the opportunities of life on the edge. Dev Biol 275, 1-11.

Hubner, N.C., Nguyen, L.N., Hornig, N.C., Stunnenberg, H.G., 2015. A quantitative proteomics tool to identify DNA-protein interactions in primary cells or blood. J Proteome Res 14, 1315-1329.

Iimura, T., Denans, N., Pourquie, O., 2009. Establishment of Hox vertebral identities in the embryonic spine precursors. Curr Top Dev Biol 88, 201-234.

Isaacs, H.V., Pownall, M.E., Slack, J.M., 1994. eFGF regulates Xbra expression during Xenopus gastrulation. EMBO J 13, 4469-4481.

Jandzik, D., Garnett, A.T., Square, T.A., Cattell, M.V., Yu, J.K., Medeiros, D.M., 2015. Evolution of the new vertebrate head by co-option of an ancient chordate skeletal tissue. Nature 518, 534-537.

Jessen, K.R., Mirsky, R., 2005. The origin and development of glial cells in peripheral nerves. Nat Rev Neurosci 6, 671-682.

Jiang, M., Stanke, J., Lahti, J.M., 2011. The connections between neural crest development and neuroblastoma. Curr Top Dev Biol 94, 77-127.

Kasparek, P., Krausova, M., Haneckova, R., Kriz, V., Zbodakova, O., Korinek, V., Sedlacek, R., 2014. Efficient gene targeting of the Rosa26 locus in mouse zygotes using TALE nucleases. FEBS Lett 588, 3982-3988.

Keenan, I.D., Sharrard, R.M., Isaacs, H.V., 2006. FGF signal transduction and the regulation of Cdx gene expression. Dev Biol 299, 478-488.

Kiecker, C., Niehrs, C., 2001. A morphogen gradient of Wnt/beta-catenin signalling regulates anteroposterior neural patterning in Xenopus. Development 128, 4189-4201.

Kim, H., Kang, K., Ekram, M.B., Roh, T.Y., Kim, J., 2011. Aebp2 as an epigenetic regulator for neural crest cells. PLoS One 6, e25174.

Kishi, M., Mizuseki, K., Sasai, N., Yamazaki, H., Shiota, K., Nakanishi, S., Sasai, Y., 2000. Requirement of Sox2-mediated signaling for differentiation of early Xenopus neuroectoderm. Development 127, 791-800.

Kobayashi, K., Jakt, L.M., Nishikawa, S.I., 2013. Epigenetic regulation of the neuroblastoma genes, Arid3b and Mycn. Oncogene 32, 2640-2648.

Kong, Q., Hai, T., Ma, J., Huang, T., Jiang, D., Xie, B., Wu, M., Wang, J., Song, Y., Wang, Y., 2014. Rosa26 locus supports tissue-specific promoter driving transgene expression specifically in pig.

Koo, T., Lee, J., Kim, J.S., 2015. Measuring and Reducing Off-Target Activities of Programmable Nucleases Including CRISPR-Cas9. Molecules and cells 38, 475-481.

Kranz, A.L., Eils, R., Konig, R., 2011. Enhancers regulate progression of development in mammalian cells. Nucleic Acids Res 39, 8689-8702.

Krispin, S., Nitzan, E., Kassem, Y., Kalcheim, C., 2010. Evidence for a dynamic spatiotemporal fate map and early fate restrictions of premigratory avian neural crest. Development 137, 585-595.

Kudoh, T., Wilson, S.W., Dawid, I.B., 2002. Distinct roles for Fgf, Wnt and retinoic acid in posteriorizing the neural ectoderm. Development 129, 4335-4346.

Levine, M., 2010. Transcriptional enhancers in animal development and evolution. Curr Biol 20, R754-763.

Lewis, A.E., Vasudevan, H.N., O'Neill, A.K., Soriano, P., Bush, J.O., 2013. The widely used Wnt1-Cre transgene causes developmental phenotypes by ectopic activation of Wnt signaling. Dev Biol 379, 229-234.

Li, J., Chen, F., Epstein, J.A., 2000. Neural crest expression of Cre recombinase directed by the proximal Pax3 promoter in transgenic mice. Genesis 26, 162-164.

Li, J., Liu, K.C., Jin, F., Lu, M.M., Epstein, J.A., 1999. Transgenic rescue of congenital heart disease and spina bifida in Splotch mice. Development 126, 2495-2503.

Li, J., Shi, Y., Sun, J., Zhang, Y., Mao, B., 2011. Xenopus reduced folate carrier regulates neural crest development epigenetically. PLoS One 6, e27198.

Limaye, A., Hall, B., Kulkarni, A.B., 2009. Manipulation of mouse embryonic stem cells for knockout mouse production. Current protocols in cell biology / editorial board, Juan S. Bonifacino ... [et al.] Chapter 19, Unit 19 13 19 13 11-24.

Little, S.C., Mullins, M.C., 2006. Extracellular modulation of BMP activity in patterning the dorsoventral axis. Birth Defects Res C Embryo Today 78, 224-242.

Liu, Y., Xiao, A., 2011. Epigenetic regulation in neural crest development. Birth Defects Res A Clin Mol Teratol 91, 788-796.

Lodish, H., Berk, A., Zipursky, S.L., al., e., 2000. Recombination between Homologous DNA Sites. Molecular Cell Biology. 4th edition., Section 12.15 Available from: http://www.ncbi.nlm.nih.gov/books/NBK21725/.

Loh, K.M., Lim, B., 2011. A precarious balance: pluripotency factors as lineage specifiers. Cell stem cell 8, 363-369.

Lohnes, D., 2003. The Cdx1 homeodomain protein: an integrator of posterior signaling in the mouse. Bioessays 25, 971-980.

Lynch, V.J., Wagner, G.P., 2008. Resurrecting the role of transcription factor change in developmental evolution. Evolution; international journal of organic evolution 62, 2131-2154.

MacDonald, B.T., Tamai, K., He, X., 2009. Wnt/beta-catenin signaling: components, mechanisms, and diseases. Dev Cell 17, 9-26.

Magli, A., Schnettler, E., Rinaldi, F., Bremer, P., Perlingeiro, R.C., 2013. Functional dissection of Pax3 in paraxial mesoderm development and myogenesis. Stem Cells 31, 59-70.

Marchant, L., Linker, C., Ruiz, P., Guerrero, N., Mayor, R., 1998. The inductive properties of mesoderm suggest that the neural crest cells are specified by a BMP gradient. Dev Biol 198, 319-329.

Marieb, E.N., Hoehn, K., 2007. Human anatomy & physiology, 7th ed. Pearson Benjamin Cummings, San Francisco.

Marletaz, F., Maeso, I., Faas, L., Isaacs, H.V., Holland, P.W., 2015. Cdx ParaHox genes acquired distinct developmental roles after gene duplication in vertebrate evolution. BMC biology 13, 56.

Marom, K., Shapira, E., Fainsod, A., 1997. The chicken caudal genes establish an anterior-posterior gradient by partially overlapping temporal and spatial patterns of expression. Mech Dev 64, 41-52.

Mason, I., 2007. Initiation to end point: the multiple roles of fibroblast growth factors in neural development. Nat Rev Neurosci 8, 583-596.

Mayor, R., Theveneau, E., 2013. The neural crest. Development 140, 2247-2251.

McBurney, M.W., 1993. P19 embryonal carcinoma cells. Int J Dev Biol 37, 135-140.

McGrew, L.L., Hoppler, S., Moon, R.T., 1997. Wnt and FGF pathways cooperatively pattern anteroposterior neural ectoderm in Xenopus. Mech Dev 69, 105-114.

McKinney, M.C., Fukatsu, K., Morrison, J., McLennan, R., Bronner, M.E., Kulesa, P.M., 2013. Evidence for dynamic rearrangements but lack of fate or position restrictions in premigratory avian trunk neural crest. Development 140, 820-830.

Meulemans, D., Bronner-Fraser, M., 2002. Amphioxus and lamprey AP-2 genes: implications for neural crest evolution and migration patterns. Development 129, 4953-4962.

Meulemans, D., Bronner-Fraser, M., 2004. Gene-regulatory interactions in neural crest evolution and development. Dev Cell 7, 291-299.

Meyer, B.I., Gruss, P., 1993. Mouse Cdx-1 expression during gastrulation. Development 117, 191-203.

Meyer, M., de Angelis, M.H., Wurst, W., Kuhn, R., 2010. Gene targeting by homologous recombination in mouse zygotes mediated by zinc-finger nucleases. Proc Natl Acad Sci U S A 107, 15022-15026.

Milewski, R.C., Chi, N.C., Li, J., Brown, C., Lu, M.M., Epstein, J.A., 2004. Identification of minimal enhancer elements sufficient for Pax3 expression in neural crest and implication of Tead2 as a regulator of Pax3. Development 131, 829-837.

Misra, R.P., Duncan, S.A., 2002. Gene targeting in the mouse: advances in introduction of transgenes into the genome by homologous recombination. Endocrine 19, 229-238.

Mlodzik, M., Fjose, A., Gehring, W.J., 1985. Isolation of caudal, a Drosophila homeo box-containing gene with maternal expression, whose transcripts form a concentration gradient at the pre-blastoderm stage. EMBO J 4, 2961-2969.

Moase, C.E., Trasler, D.G., 1992. Splotch locus mouse mutants: models for neural tube defects and Waardenburg syndrome type I in humans. J Med Genet 29, 145-151.

Monsoro-Burq, A.H., Wang, E., Harland, R., 2005. Msx1 and Pax3 cooperate to mediate FGF8 and WNT signals during Xenopus neural crest induction. Dev Cell 8, 167-178.

Mulley, J.F., Chiu, C.H., Holland, P.W., 2006. Breakup of a homeobox cluster after genome duplication in teleosts. Proc Natl Acad Sci U S A 103, 10369-10372.

Mulley, J.F., Holland, P.W., 2010. Parallel retention of Pdx2 genes in cartilaginous fish and coelacanths. Molecular biology and evolution 27, 2386-2391.

Niehrs, C., 2004. Regionally specific induction by the Spemann-Mangold organizer. Nat Rev Genet 5, 425-434.

Nyabi, O., Naessens, M., Haigh, K., Gembarska, A., Goossens, S., Maetens, M., De Clercq, S., Drogat, B., Haenebalcke, L., Bartunkova, S., De Vos, I., De Craene, B., Karimi, M., Berx, G., Nagy, A., Hilson, P., Marine, J.C., Haigh, J.J., 2009. Efficient mouse transgenesis using Gateway-compatible ROSA26 locus targeting vectors and F1 hybrid ES cells. Nucleic Acids Res 37, e55.

Okuda, Y., Ogura, E., Kondoh, H., Kamachi, Y., 2010. B1 SOX coordinate cell specification with patterning and morphogenesis in the early zebrafish embryo. PLoS Genet 6, e1000936.

Oliveri, P., Davidson, E.H., McClay, D.R., 2003. Activation of pmar1 controls specification of micromeres in the sea urchin embryo. Dev Biol 258, 32-43.

Owens, J.B., Mauro, D., Stoytchev, I., Bhakta, M.S., Kim, M.S., Segal, D.J., Moisyadi, S., 2013. Transcription activator like effector (TALE)-directed piggyBac transposition in human cells. Nucleic Acids Res 41, 9197-9207.

Papanayotou, C., Mey, A., Birot, A.M., Saka, Y., Boast, S., Smith, J.C., Samarut, J., Stern, C.D., 2008. A mechanism regulating the onset of Sox2 expression in the embryonic neural plate. PLoS Biol 6, e2.

Paratore, C., Eichenberger, C., Suter, U., Sommer, L., 2002. Sox10 haploinsufficiency affects maintenance of progenitor cells in a mouse model of Hirschsprung disease. Hum Mol Genet 11, 3075-3085.

Pares-Matos, E.I., 2013. Electrophoretic mobility-shift and super-shift assays for studies and characterization of protein-DNA complexes. Methods Mol Biol 977, 159-167.

Patten, I., Placzek, M., 2002. Opponent activities of Shh and BMP signaling during floor plate induction in vivo. Curr Biol 12, 47-52.

Peinado, H., Portillo, F., Cano, A., 2004. Transcriptional regulation of cadherins during development and carcinogenesis. Int J Dev Biol 48, 365-375.

Pennacchio, L.A., Ahituv, N., Moses, A.M., Prabhakar, S., Nobrega, M.A., Shoukry, M., Minovitsky, S., Dubchak, I., Holt, A., Lewis, K.D., Plajzer-Frick, I., Akiyama, J., De Val, S., Afzal, V., Black, B.L., Couronne, O., Eisen, M.B., Visel, A., Rubin, E.M., 2006. In vivo enhancer analysis of human conserved non-coding sequences. Nature 444, 499-502.

Pennacchio, L.A., Bickmore, W., Dean, A., Nobrega, M.A., Bejerano, G., 2013. Enhancers: five essential questions. Nat Rev Genet 14, 288-295.

Pennisi, E., 2008. Evolutionary biology. Deciphering the genetics of evolution. Science 321, 760-763.

Perris, R., 1997. The extracellular matrix in neural crest-cell migration. Trends Neurosci 20, 23-31.

Petersen, C.P., Reddien, P.W., 2009. Wnt signaling and the polarity of the primary body axis. Cell 139, 1056-1068.

Pevny, L.H., Nicolis, S.K., 2010. Sox2 roles in neural stem cells. Int J Biochem Cell Biol 42, 421-424.

Pilon, N., Oh, K., Sylvestre, J.R., Bouchard, N., Savory, J., Lohnes, D., 2006. Cdx4 is a direct target of the canonical Wnt pathway. Dev Biol 289, 55-63.

Pilon, N., Oh, K., Sylvestre, J.R., Savory, J.G., Lohnes, D., 2007. Wnt signaling is a key mediator of Cdx1 expression in vivo. Development 134, 2315-2323.

Pitkin, R.M., 2007. Folate and neural tube defects. Am J Clin Nutr 85, 285S-288S.

Potterf, S.B., Mollaaghababa, R., Hou, L., Southard-Smith, E.M., Hornyak, T.J., Arnheiter, H., Pavan, W.J., 2001. Analysis of SOX10 function in neural crest-derived melanocyte development: SOX10-dependent transcriptional control of dopachrome tautomerase. Developmental biology 237, 245-257.

Prinos, P., Joseph, S., Oh, K., Meyer, B.I., Gruss, P., Lohnes, D., 2001. Multiple pathways governing Cdx1 expression during murine development. Dev Biol 239, 257-269.

Pruitt, S.C., Bussman, A., Maslov, A.Y., Natoli, T.A., Heinaman, R., 2004. Hox/Pbx and Brn binding sites mediate Pax3 expression in vitro and in vivo. Gene Expr Patterns 4, 671-685.

Puliti, A., Prehu, M.O., Simon-Chazottes, D., Ferkdadji, L., Peuchmaur, M., Goossens, M., Guenet, J.L., 1995. A high-resolution genetic map of mouse chromosome 15 encompassing the Dominant megacolon (Dom) locus. Mamm Genome 6, 763-768.

Purves, D., Williams, S.M., 2001. Neuroscience, 2nd ed. Sinauer Associates, Sunderland, Mass.

Rada-Iglesias, A., Bajpai, R., Prescott, S., Brugmann, S.A., Swigut, T., Wysocka, J., 2012. Epigenomic annotation of enhancers predicts transcriptional regulators of human neural crest. Cell Stem Cell 11, 633-648.

Rada-Iglesias, A., Prescott, S.L., Wysocka, J., 2013. Human genetic variation within neural crest enhancers: molecular and phenotypic implications. Philos Trans R Soc Lond B Biol Sci 368, 20120360.

Rex, M., Orme, A., Uwanogho, D., Tointon, K., Wigmore, P.M., Sharpe, P.T., Scotting, P.J., 1997. Dynamic expression of chicken Sox2 and Sox3 genes in ectoderm induced to form neural tissue. Dev Dyn 209, 323-332.

Rhead, B., Karolchik, D., Kuhn, R.M., Hinrichs, A.S., Zweig, A.S., Fujita, P.A., Diekhans, M., Smith, K.E., Rosenbloom, K.R., Raney, B.J., Pohl, A., Pheasant, M., Meyer, L.R., Learned, K., Hsu, F., Hillman-Jackson, J., Harte, R.A., Giardine, B., Dreszer, T.R., Clawson, H., Barber, G.P., Haussler, D., Kent, W.J., 2010. The UCSC Genome Browser database: update 2010. Nucleic Acids Res 38, D613-619.

Russel, W., 1947. Splotch, a new mutation in the house mouse. Mus musculus.

Sanchez-Ferras, O., Bernas, G., Laberge-Perrault, E., Pilon, N., 2014. Induction and dorsal restriction of Paired-box 3 (Pax3) gene expression in the caudal neuroectoderm is mediated by integration of multiple pathways on a short neural crest enhancer. Biochim Biophys Acta 1839, 546-558.

Sanchez-Ferras, O., Coutaud, B., Djavanbakht Samani, T., Tremblay, I., Souchkova, O., Pilon, N., 2012. Caudal-related homeobox (Cdx) protein-dependent integration of canonical Wnt signaling on paired-box 3 (Pax3) neural crest enhancer. J Biol Chem 287, 16623-16635.

Santarius, T., Shipley, J., Brewer, D., Stratton, M.R., Cooper, C.S., 2010. A census of amplified and overexpressed human cancer genes. Nature Reviews Cancer 10, 59-64.

Sato, T., Sakuma, T., Yokonishi, T., Katagiri, K., Kamimura, S., Ogonuki, N., Ogura, A., Yamamoto, T., Ogawa, T., 2015. Genome Editing in Mouse Spermatogonial Stem Cell Lines Using TALEN and Double-Nicking CRISPR/Cas9. Stem cell reports 5, 75-82.

Savory, J.G., Bouchard, N., Pierre, V., Rijli, F.M., De Repentigny, Y., Kothary, R., Lohnes, D., 2009a. Cdx2 regulation of posterior development through non-Hox targets. Development 136, 4099-4110.

Savory, J.G., Mansfield, M., Rijli, F.M., Lohnes, D., 2011a. Cdx mediates neural tube closure through transcriptional regulation of the planar cell polarity gene Ptk7. Development 138, 1361-1370.

Savory, J.G., Mansfield, M., St Louis, C., Lohnes, D., 2011b. Cdx4 is a Cdx2 target gene. Mech Dev 128, 41-48.

Savory, J.G., Pilon, N., Grainger, S., Sylvestre, J.R., Beland, M., Houle, M., Oh, K., Lohnes, D., 2009b. Cdx1 and Cdx2 are functionally equivalent in vertebral patterning. Dev Biol 330, 114-122.

Schulte, J.H., Lindner, S., Bohrer, A., Maurer, J., De Preter, K., Lefever, S., Heukamp, L., Schulte, S., Molenaar, J., Versteeg, R., Thor, T., Kunkele, A., Vandesompele, J., Speleman, F., Schorle, H., Eggert, A., Schramm, A., 2013. MYCN and ALKF1174L are sufficient to drive neuroblastoma development from neural crest progenitor cells. Oncogene 32, 1059-1065.

Sever, L.E., 1995. Looking for causes of neural tube defects: where does the environment fit in? Environ Health Perspect 103 Suppl 6, 165-171.

Shakhova, O., Zingg, D., Schaefer, S.M., Hari, L., Civenni, G., Blunschi, J., Claudinot, S., Okoniewski, M., Beermann, F., Mihic-Probst, D., Moch, H., Wegner, M., Dummer, R., Barrandon, Y., Cinelli, P., Sommer, L., 2012. Sox10 promotes the formation and maintenance of giant congenital naevi and melanoma. Nat Cell Biol 14, 882-890.

Shastry, B., 1995. Overexpression of genes in health and sickness. A bird's eye view. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology 112, 1-13.

Shimizu, T., Bae, Y.K., Muraoka, O., Hibi, M., 2005. Interaction of Wnt and caudal-related genes in zebrafish posterior body formation. Dev Biol 279, 125-141.

Shirley, S.H., Greene, V.R., Duncan, L.M., Torres Cabala, C.A., Grimm, E.A., Kusewitt, D.F., 2012. Slug expression during melanoma progression. Am J Pathol 180, 2479-2489.

Silberg, D.G., Swain, G.P., Suh, E.R., Traber, P.G., 2000. Cdx1 and cdx2 expression during intestinal development. Gastroenterology 119, 961-971.

Simoes-Costa, M., Bronner, M.E., 2013. Insights into neural crest development and evolution from genomic analysis. Genome Res 23, 1069-1080.

Simoes-Costa, M., Bronner, M.E., 2015. Establishing neural crest identity: a gene regulatory recipe. Development 142, 242-257.

Smith, J.L., Schoenwolf, G.C., 1997. Neurulation: coming to closure. Trends Neurosci 20, 510-517.

Smith, W.C., Harland, R.M., 1992. Expression cloning of noggin, a new dorsalizing factor localized to the Spemann organizer in Xenopus embryos. Cell 70, 829-840.

Soriano, P., 1999. Generalized lacZ expression with the ROSA26 Cre reporter strain. Nat Genet 21, 70-71.

Southard-Smith, E.M., Kos, L., Pavan, W.J., 1998. Sox10 mutation disrupts neural crest development in Dom Hirschsprung mouse model. Nat Genet 18, 60-64.

Srinivas, S., Watanabe, T., Lin, C.-S., William, C.M., Tanabe, Y., Jessell, T.M., Costantini, F., 2001. Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. BMC Developmental Biology 1.

Stottmann, R.W., Berrong, M., Matta, K., Choi, M., Klingensmith, J., 2006. The BMP antagonist Noggin promotes cranial and spinal neurulation by distinct mechanisms. Dev Biol 295, 647-663.

Strumpf, D., Mao, C.A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., Rossant, J., 2005. Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. Development 132, 2093-2102.

Stuhlmiller, T.J., Garcia-Castro, M.I., 2012. Current perspectives of the signaling pathways directing neural crest induction. Cell Mol Life Sci 69, 3715-3737.

Subramanian, V., Meyer, B.I., Gruss, P., 1995. Disruption of the murine homeobox gene Cdx1 affects axial skeletal identities by altering the mesodermal expression domains of Hox genes. Cell 83, 641-653.

Takemoto, T., Uchikawa, M., Kamachi, Y., Kondoh, H., 2006. Convergence of Wnt and FGF signals in the genesis of posterior neural plate through activation of the Sox2 enhancer N-1. Development 133, 297-306.

Tam, P.P., Behringer, R.R., 1997. Mouse gastrulation: the formation of a mammalian body plan. Mech Dev 68, 3-25.

Thomson, M., Liu, S.J., Zou, L.-N., Smith, Z., Meissner, A., Ramanathan, S., 2011. Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. Cell 145, 875-889.

Trainor, P., 2013. Neural crest cells: evolution, development and disease. Academic Press.

Trainor, P.A., 2010. Craniofacial birth defects: The role of neural crest cells in the etiology and pathogenesis of Treacher Collins syndrome and the potential for prevention. Am J Med Genet A 152A, 2984-2994.

Tronche, F., Kellendonk, C., Kretz, O., Gass, P., Anlag, K., Orban, P.C., Bock, R., Klein, R., Schutz, G., 1999. Disruption of the glucocorticoid receptor gene in the nervous system results in reduced anxiety. Nat Genet 23, 99-103.

Tuazon, F.B., Mullins, M.C., 2015. Temporally coordinated signals progressively pattern the anteroposterior and dorsoventral body axes. Semin Cell Dev Biol.

Uchikawa, M., Yoshida, M., Iwafuchi-Doi, M., Matsuda, K., Ishida, Y., Takemoto, T., Kondoh, H., 2011. B1 and B2 Sox gene expression during neural plate development in chicken and mouse embryos: universal versus species-dependent features. Dev Growth Differ 53, 761-771.

van den Akker, E., Forlani, S., Chawengsaksophak, K., de Graaff, W., Beck, F., Meyer, B.I., Deschamps, J., 2002. Cdx1 and Cdx2 have overlapping functions in anteroposterior patterning and posterior axis elongation. Development 129, 2181-2193.

van Nes, J., de Graaff, W., Lebrin, F., Gerhard, M., Beck, F., Deschamps, J., 2006. The Cdx4 mutation affects axial development and reveals an essential role of Cdx genes in the ontogenesis of the placental labyrinth in mice. Development 133, 419-428.

Van Otterloo, E., Cornell, R.A., Medeiros, D.M., Garnett, A.T., 2013. Gene regulatory evolution and the origin of macroevolutionary novelties: insights from the neural crest. Genesis 51, 457-470.

Wakamatsu, Y., Watanabe, Y., Nakamura, H., Kondoh, H., 1997. Regulation of the neural crest cell fate by N-myc: promotion of ventral migration and neuronal differentiation. Development 124, 1953-1962.

Wang, W.C., Shashikant, C.S., 2007. Evidence for positive and negative regulation of the mouse Cdx2 gene. J Exp Zool B Mol Dev Evol 308, 308-321.

Weiss, M.B., Abel, E.V., Mayberry, M.M., Basile, K.J., Berger, A.C., Aplin, A.E., 2012. TWIST1 is an ERK1/2 effector that promotes invasion and regulates MMP-1 expression in human melanoma cells. Cancer Res 72, 6382-6392.

Weiss, W.A., Aldape, K., Mohapatra, G., Feuerstein, B.G., Bishop, J.M., 1997. Targeted expression of MYCN causes neuroblastoma in transgenic mice. EMBO J 16, 2985-2995.

Weston, J.A., Butler, S.L., 1966. Temporal factors affecting localization of neural crest cells in tbe chicken embryo. Dev Biol 14, 246-266.

Wittkopp, P.J., Kalay, G., 2012. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. Nature Reviews Genetics 13, 59-69.

Wray, G.A., 2007. The evolutionary significance of cis-regulatory mutations. Nat Rev Genet 8, 206-216.

Xu, F., Li, H., Jin, T., 1999. Cell type-specific autoregulation of the Caudal-related homeobox gene Cdx-2/3. J Biol Chem 274, 34310-34316.

Yamaguchi, Y., Miura, M., 2013. How to form and close the brain: insight into the mechanism of cranial neural tube closure in mammals. Cell Mol Life Sci 70, 3171-3186.

Yang, J., Mani, S.A., Donaher, J.L., Ramaswamy, S., Itzykson, R.A., Come, C., Savagner, P., Gitelman, I., Richardson, A., Weinberg, R.A., 2004. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. Cell 117, 927-939.

Ybot-Gonzalez, P., Cogram, P., Gerrelli, D., Copp, A.J., 2002. Sonic hedgehog and the molecular regulation of mouse neural tube closure. Development 129, 2507-2517.

Ybot-Gonzalez, P., Gaston-Massuet, C., Girdler, G., Klingensmith, J., Arkell, R., Greene, N.D., Copp, A.J., 2007. Neural plate morphogenesis during mouse neurulation is regulated by antagonism of Bmp signalling. Development 134, 3203-3211.

Zambrowicz, B.P., Imamoto, A., Fiering, S., Herzenberg, L.A., Kerr, W.G., Soriano, P., 1997. Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. Proc Natl Acad Sci U S A 94, 3789-3794.

Zhang, S., 2014. Sox2, a key factor in the regulation of pluripotency and neural differentiation. World Journal of Stem Cells 6, 305.

Zhao, T., Gan, Q., Stokes, A., Lassiter, R.N., Wang, Y., Chan, J., Han, J.X., Pleasure, D.E., Epstein, J.A., Zhou, C.J., 2014. beta-catenin regulates Pax3 and Cdx2 for caudal neural tube closure and elongation. Development 141, 148-157.