Regulation of Axial Elongation by *Cdx*

Yalun Zhu

Thesis submitted to the University of Ottawa

in partial Fulfillment of the requirements for the

MSc in Cellular and Molecular Medicine

Department of Cellular and Molecular Medicine

Faculty of Medicine

University of Ottawa

AUTHORIZATION

Figure 3 (Lohnes 2003). Reproduced with permission from the BioEssays publishing group

Figure 4 (Böhmer, C., Werneburg, I. 2017). Reproduced with permission from the Scientific Reports publishing group.

ABSTRACT

During mouse development, the primordia of the posterior body including the trunk and tail tissues of the embryo forms largely from a bipotential cell population that resides in the posterior growth zone in vertebrate embryos. This bipotential cell population contains neuromesodermal progenitors (NMP) which are found in the tail bud which replaces the primitive streak after gastrulation and contributes to axial elongation by the formation of both the spinal cord and paraxial mesoderm derivatives. The three vertebrate *Cdx* genes, *Cdx1, Cdx2* and *Cdx4,* encode transcription factors that play important roles in axial elongation since the triple *Cdx* mutant embryos fail to generate any tissue posterior to the occipital primordia. A comparison of *Cdx* mutant phenotypes suggests that Cdx2 is the most important contributor to axial elongation since *Cdx2* heterozygous mutants exhibit foreshortened tails and *Cdx2* conditional mutants exhibit axial truncation and complete loss of tail bud structures. Cdx2 target genes, such as *Wnt3a, Cyp26a1* and *T,* are also essential for axial elongation. *Cdx1* null mutants are viable and exhibit homeosis of cervical and anterior thoracic vertebrae, while *Cdx4* null mutants are phenotypically normal. In addition, it has been shown that simultaneous loss of multiple copies of *Cdx* alleles disrupts axial elongation more severely than each single mutation which suggests there is overlapping function among the *Cdx* family. The genetic network underlying regulation of axial elongation by the *Cdx* family is not fully understood due in part to this functional overlap. In this thesis, I employed a conditional Cre-loxP system to derive conditional mutants lacking all Cdx functions. Additionally, *Pax2-GFP* transgenic mice where GFP is expressed under the control of *Pax2* locus were used to enrich tail bud NMP cells for RNA-seq and ChIP-seq analysis for Cdx2. Using this approach, I revealed new target genes and pathways that are regulated by Cdx members and likely involved in axial elongation.

Table of Contents

TABLE OF FIGURES

LIST OF TABLE

ABBREVIATIONS

Abbreviation	Definition
ADH	Alcohol Dehydrogenases
AP	Anterior-Posterior
APC	Adenomatous Polyposis Coli
AVE	Anterior Visceral Endoderm
BP	Base Pair
CDRE	Cdx Response Element
CDX	Caudal-Related Homeobox
ChIP	Chromatin Immunoprecipitation
ChIP-seq	ChIP-sequencing
CK1a	Casein Kinase 1a
CLE	Caudal-Lateral Epiblast
CNH	Chordo-Neural Hinge
CRE	*cis*-Regulatory Elements
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CTCF	CCCTC-binding factor
CYP26A1	Cytochrome P450, Family 26, Subfamily A, Polypeptide 1
DKO	Double Knockout
DNA	Deoxyribonucleic Acid
Dsh	Dishevelled
E	Embryonic Day

EDTA	Ethylenediaminetetraacetic Acid
EMT	Epithelial-Mesenchymal Transition
ER^T	Estrogen Receptor Ligand Domain modified to respond to Tamoxifen
ESC	Embryonic Stem Cell
ExE	Extraembryonic Ectoderm
FACS	Fluorescence Activated Cell Sorting
FGF	Fibroblast Growth Factor
FGFR	FGF Tyrosine Kinase Receptor
Flox	Flanked by LoxP
FZ	Frizzled
GFP	Green Fluorescent Protein
GO	Gene Ontology
GSK3	Glycogen Synthase Kinase 3
Hi-C seq	High-throughput Chromosome Conformation Capture sequencing
ICM	Inner Cell Mass
ISH	*In Situ* Hybridization
JNK	Jun Kinase
LEF	Lymphoid Enhancer Binding Factor
LoxP	Locus of Crossover in P1
LRP5/6	Low-Density-Lipoprotein-Related Protein 5/6
MO	Morpholino-Mediated
NMP	Neuro-Mesodermal Progenitor
NR2F1	Nuclear Receptor Subfamily 2 Group F Member 1

NSB	Node-Streak Border
PBS	Phosphate-Buffered Saline
PCA	Principal Component Analysis
PCP	Planar Cell Polarity Pathway
PCR	Polymerase Chain Reaction
PFA	Paraformaldehyde
PM	Paraxial Mesoderm
PP2A	Protein Phosphatase 2A
PS	Primitive Streak
PSM	Presomitic Mesoderm
qPCR	Quantitative Polymerase Chain Reaction
RA	Retinoic Acid
RALDH	Retinaldehyde Dehydrogenase
RAR	RA Receptor
RDH	Retinol Dehydrogenase
RIN	RNA Integrity Number
RNA	Ribonucleic Acid
RNA-seq	RNA-sequencing
RT-PCR	Reverse-Transcriptase Polymerase Chain Reaction
RXR	Retinoid X Receptors
scRNA-seq	Single Cell RNA-seq
SMC	Structural Maintenance of Chromosomes
SWI-SNF	Switch/Sucrose-Non-Fermentable
T	Brachyury

TAD	Topologically Associated Domain
TB	Tail Bud
TBST	Tris-Buffered Saline Tween-20
TCF	T-Cell Factor
TSS	Transcriptional Start Site
VE	Visceral Endoderm
WNT	Wingless-Related
WT	Wild Type

ACKNOWLEDGEMENTS

Throughout the writing of this thesis, I have received a great deal of support and assistance from a number of individuals directly or indirectly in shaping up my academic career.

First and foremost, I would like to express my deepest appreciation to my mentor Dr. David Lohnes for introducing me to the world of research and science. It was due to his valuable guidance, cheerful enthusiasm and full support that I was able to complete my research in a respectable manner. I am also extremely grateful to him for allowing me some extra time to finish my experimental work. It was because of him that I learned how to ask questions and find answers, the importance of writing logically and concisely, and I acquired inspiration every time we talk. I will never forget the contributions that you made to my success.

I would also like to thank my thesis advisory committee members Dr. Marie-Andrée Akimenko and Dr. Nadine Wiper-Bergeron for their guidance and challenges throughout my degree. Our meetings always stimulated new ideas and helped me move my project along.

I also had great pleasure of working with every member of the Lohnes lab, who have helped me throughout my master's degree. Tanya, your support and encouragement have helped me through the frustration of my life, and I will always remember the great times we had in and outside of the lab as well. Brad, you are the first one who I talked to in the lab, and you always tried to keep lab organized and clean which makes my life a lot easier. Sanzida, I really appreciate the advice that you gave me and the experiences that you shared with me, which saved me a lot of time.

Wei, thank you for helping me overcome the fear of touching mice and your great help on mouse colony. Simon, thank you for your selfless sacrifices that helped others, and I will never forget the one time you helped me when I hit the stone in curling. Faiyaz, my late-night lab bro, thank you for always being there for me. Adriana, also known as lab grandma, thank you for warm words and your delicious tiramisu.xiii

Finally, I am indebted to my parents, especially to my mom, for being close friends and supporting me even during hard times. The present work would not have been possible without this network of love and support. Thank you all!

Chapter 1

INTRODUCTION

Early mouse development

After fertilization (E0), the single-cell mouse embryo undergoes a series of symmetrical cleavages to produce a 16-cell morula which begins the cavitation process that results in the blastocyst at E3.5 (Fleming, 1987). The blastocyst is polarized such that the inner cell mass (ICM) and adjacent trophectoderm form the embryonic pole of blastocyst whereas the mural trophectoderm and the blastocoel cavity form the abembryonic pole. At E4.5, a group of ICM cells differentiate into primitive endoderm and the remaining ICM cells become the primitive ectoderm, or epiblast, cells. The epiblast cells, the trophectoderm and primitive endoderm are the progenitors for all tissues of the embryo and the extra-embryonic tissues that support embryo development. The trophectoderm cells contribute to trophoblast cells and extraembryonic ectoderm that make up the majority of the fetal part of the placenta. The primitive endoderm cells develop into extraembryonic endoderm layers of visceral endoderm (VE) and parietal endoderm. The epiblast cells give rise to all three germ layers of the embryo proper and to the extra-embryonic mesoderm of the visceral yolk sac, allantois and the hematopoietic precursors (Gardner, 1978).

The series of symmetrical cleavages leading to the blastocyst all take place in the oviduct (Nagy et al., 2003). Then at E4.5, the blastocyst implants in the uterus mediated by trophectoderm

where the epiblast cells start to proliferate and differentiate rapidly and the embryo proper undergoes complex morphological processes leading to a key developmental stage, gastrulation, starting at E6.0. At this point, the embryo proper is a cup-shaped epiblast and the gastrulation commences with formation of the primitive streak. The location of the primitive streak is considered as the posterior side of the embryo (Arnold and Robertson, 2009; Beddington and Robertson, 1999). The epiblast cells ingress through the primitive streak and undergo epithelial-mesenchymal transformation (Acloque et al., 2009) to form nascent mesoderm and definitive endoderm and the non-ingressing epiblast cells form the ectoderm (Tam and Behringer, 1997); all the fetal tissues will develop from these three germ layers. The ectoderm gives rise to the nervous system, epidermis and neural crest cells. The mesoderm gives rise to the diverse tissue types, including muscle, axial and appendicular skeleton, blood, gonads, kidney and associated ducts, and connective tissues. The endoderm gives rise to the digestive tract and internal organs.

As gastrulation progresses, starting from E6.25 the primitive streak elongates and extends towards the anterior tip of the embryo. Mesoderm leaves the primitive streak region immediately after ingression and the progressive addition of new tissue at the posterior end of the embryo results in the elongation of the body axis (Young et al., 2009). During this process, the mesodermal progenitors will further differentiate into notochord, lateral mesoderm (circulatory system, limb bud mesenchyme and walls of the digestive organs), intermediate mesoderm (urogenital system) or paraxial mesoderm (presomitic mesoderm, somites and associated connective tissue and musculature) (Wolpert, 2019).

Axial elongation

Axial elongation is dependent on the existence and maintenance of progenitors initially residing in the anteriormost part of the primitive streak (Cambray and Wilson, 2007). Clonal analysis using serial grafting in early mouse embryos has traced this progenitor cell population to the node-streak border (NSB) area, the caudal-lateral epiblast (CLE) and the chordo-neural hinge (CNH) in the tail bud that functions as a continuation of the primitive streak which migrates caudally and disappears (Cambray and Wilson, 2002). Fate mapping analyses suggest that these progenitor cell populations are bipotent and contribute descendants to both neuroectoderm and mesoderm during axial elongation to give rise to trunk and tail structures; these progenitor cells are therefore typically defined as neuro-mesodermal progenitors (NMPs) (Fig. 1) (Tzouanacou et al., 2009).

NMPs are characterised by the coexpression of transcription factors Sox2 and T/Brachyury, which are required for development of the neural and mesodermal lineages, respectively (Martin and Kimelman, 2012; Olivera-Martinez et al., 2012; Tsakiridis et al., 2014). Therefore, the development of spinal cord and paraxial mesoderm (PM) derived tissues in the trunk and tail requires the proper differentiation and maintenance of the NMPs (Cambray and Wilson, 2002; García-García et al., 2008; Mathis and Nicolas, 2000).

Genetic studies have shown critical roles for several signaling pathways and other transcription factors in the maintenance and differentiation of NMPs, and thus the elaboration of the body axes during development. Among these are the canonical Wnt (Takada et al., 1994), Retinoic Acid

(Janesick et al., 2014; Olivera-Martinez et al., 2012) and Fgf (Boulet and Capecchi, 2012; Olivera-Martinez et al., 2012) pathways and the transcription factors Cdx (Chawengsaksophak et al., 2004; Lohnes, 2003; Savory et al., 2011a; Savory et al., 2009a; van den Akker et al., 2002; van Rooijen et al., 2012) and T/Brachyury (Herrmann et al., 1990), all of which have been shown to be essential for anteroposterior (AP) axial elongation.

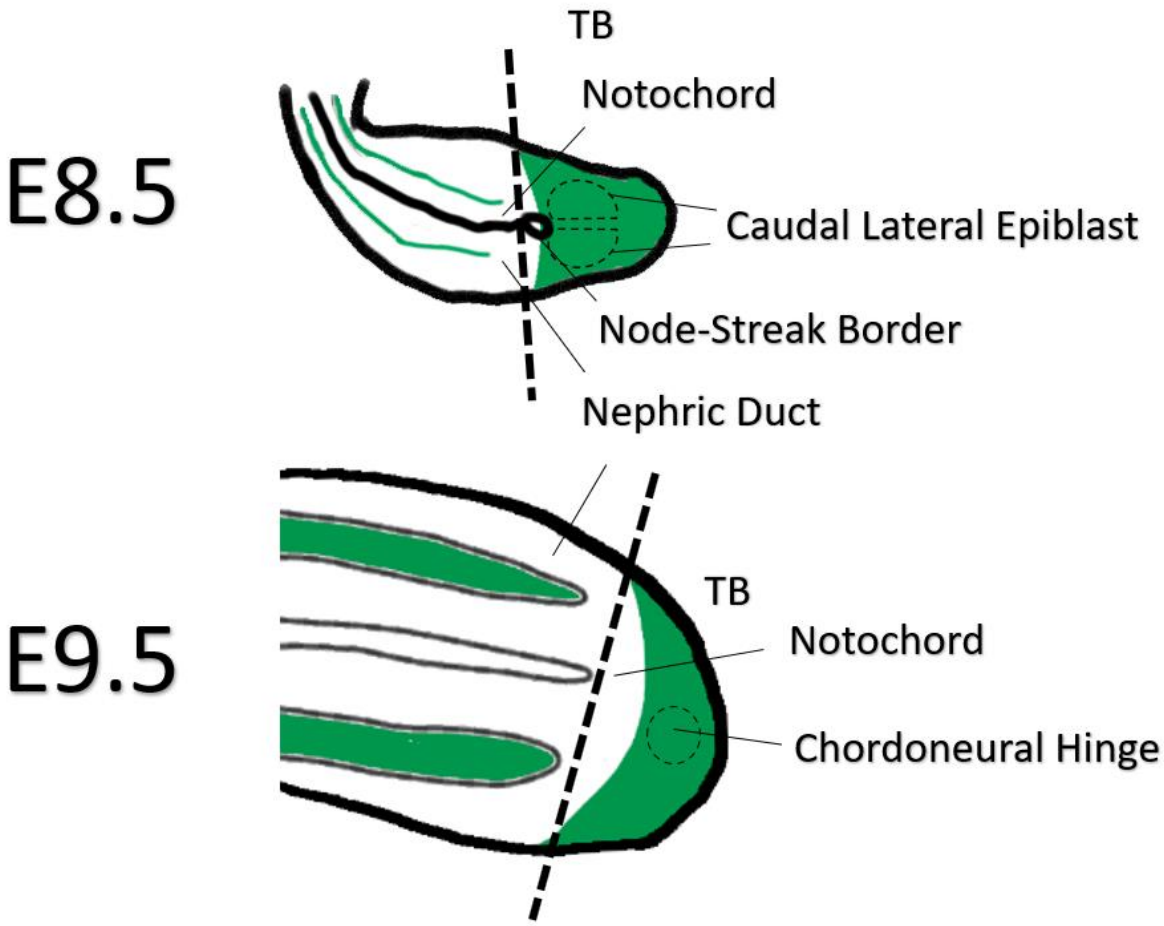


Figure 1. Neuromesodermal Progenitor Cells in the Tail bud of Mouse Embryo. Schematic representation of mouse embryos at E8.5 and E9.5. Pax2-GFP is expressed in the neuromesoderm resident regions (CLE, NSB and CNH) and nephric duct. The solid line indicates the postcranial structures of the mouse embryo which are formed continuously by addition of tissues from the neuromesodermal progenitor cells located in the tail bud, which is the posteriormost part of the embryo colored in green. The dashed straight line indicates the dissected regions.

Canonical Wnt Signaling in development

Wnt signalling is evolutionarily conserved in metazoans (Clevers and Nusse, 2012). The term Wnt is derived from the combination of the names of the *Drosophila* segment polarity gene wingless and the vertebrate homolog int-1 (Croce and McClay, 2008). It has been well documented that Wnt signalling is essential in embryonic development and impacts body axis patterning, cell fate specification, proliferation, migration and maintenance of stem cell populations during early embryogenesis as well as in adulthood (Clevers and Nusse, 2012; Hikasa and Sokol, 2013). Mutations in the members of the Wnt signalling pathway typically lead to serious developmental defects and diseases in adult life such as cancer (Clevers and Nusse, 2012).

Wnt signals are able to transduce via three distinct pathways (Korswagen, 2002): the canonical Wnt pathway, the planar cell polarity pathway and the Wnt/calcium pathway. The non-canonical pathways utilize many of the same components as the canonical pathway but with different transducing molecules (Wharton, 2003). All three pathways are triggered by binding of Wnt ligand to a heterodimeric receptor complex formed by a Frizzled (Fz) and the low-density-lipoprotein-related protein 5/6 (LRP5/6) protein, followed by the activation of the cytoplasmic effector Dishevelled (Dsh) inside the cell (van Amerongen and Nusse, 2009). The canonical Wnt pathway culminates in the stabilization and nuclear accumulation of β-catenin (Clevers and Nusse, 2012). Without Wnt signaling, cytoplasmic β-catenin is degraded by a β-catenin destruction complex, which includes Axin, adenomatous polyposis coli (APC), protein phosphatase 2A (PP2A), glycogen synthase kinase 3 (GSK3) and casein kinase 1a (CK1a) (Gordon and Nusse, 2006). Wnt ligands inhibit this destructor complex, leading to nuclear

accumulation of β-catenin and transcriptional upregulation of target genes via physical association with members of the LEF-TCF transcription factor family. The mutations in certain core components of the destruction complex, such as APC, lead to the failure of β-catenin degradation and consequently result in aberrant Wnt target gene expression that can cause a hereditary cancer syndrome called familial adenomatous polyposis in the intestinal tract (Nishisho et al., 1991).

The non-canonical planar cell polarity (PCP) pathway is also mediated by ligand activation via Fz and Dsh, but it does not involve β-catenin stabilization (Simons and Mlodzik, 2008), and requires the activation of small GTPases Rho and Rac instead (Wallingford and Habas, 2005). These in turn activate the Rock (Rho kinase) and JNK (Jun Kinase) leading to actin polymerization and affecting cell polarization and motility during gastrulation (Komiya and Habas, 2008).

The non-canonical Wnt/Ca^{2+} signaling transduces signal through Fz to stimulate intracellular Ca^{2+} release from the endoplasmic reticulum (ER) mediated by activation of Dsh via activation of G-proteins (Kohn and Moon, 2005). It is proposed that the roles of Wnt/Ca^{2+} pathway include tissue separation and cell movements during gastrulation (Komiya and Habas, 2008).

The canonical Wnt pathway is required for many developmental processes including axial elongation (Hikasa and Sokol, 2013). Wnt1, Wnt3, Wnt3a and Wnt8 are canonical Wnt ligands that function through β-catenin stabilization and transcriptional activation of Wnt target genes

(Clevers and Nusse, 2012), with Wnt3 and Wnt3a being essential to axial elongation of vertebrate embryos. *Wnt3* expression initiates in the visceral endoderm (VE) at E5.5 and is present in the posterior epiblast at about E5.75 (Rivera-Pérez and Magnuson, 2005). The anterior visceral endoderm (AVE) of *Wnt3* null mice is formed, therefore the *Wnt3* null mice have an anterior posterior axis, however the axis is not elongated due to a lack of primitive streak formation and failure to gastrulate (Liu et al., 1999). The early lethality of *Wnt3* null mouse embryos prevented analysis of later functions for this ligand. In this case, studies using a conditional *Wnt3* allele have been performed. *Wnt3* conditional mutant embryos showed primitive streak appearance and initiation of gastrulation, but embryos were delayed in development and died *in utero* at E9.5. These results indicated that *Wnt3* is necessary for continuation of axial elongation after gastrulation (Tortelote et al., 2013).

Wnt3a is expressed in the primitive streak and tail bud of the developing embryo (Takada et al., 1994). *Wnt3a* null mice have a severe axial truncation commencing at the forelimb, and a disrupted notochord, while anterior structures remain unaffected (Greco et al., 1996). *Wnt3a* is also required for the oscillating expression of Notch signalling in the presomitic mesoderm (PSM) to regulate somitogenesis (Aulehla et al., 2003). The ectopic neural tissues of *Wnt3a* null mice results from the misspecification of paraxial mesoderm since *Wnt3a* is essential for T/Brachyury expression and a paraxial mesoderm fate choice (Yoshikawa et al., 1997).

FGF signaling in axial elongation

Fibroblast growth factors (FGFs) are a family of extracellular polypeptide signaling ligands that bind to FGF tyrosine kinase receptors (FGFRs). In vertebrates, the FGF family has 22 ligands identified by their sequence homology and function (Ornitz, 2000). The FGFR family contains four members *FGFR1-4* which undergo alternative splicing in their extracellular domain resulting in receptor variants with different affinities for FGF ligands (Zhang et al., 2006). The binding of the FGF ligands to the extracellular domain of the transmembrane tyrosine kinase receptors results in the transphosphorylation of intracellular tyrosine residues and activates a MAPK cascade which eventually targets gene expression (Teven et al., 2014). Since their discovery, FGFs have been implicated in regulating crucial developmental events including formation of the primitive streak, induction and maintenance of mesoderm and neuroectoderm, and axial elongation (Dorey and Amaya, 2010).

Extensive analysis in mice has suggested that FGF ligands Fgf4 and Fgf8, and the main Fgf8 receptor Fgfr1, play a critical role in axial elongation. *Fgf4* is expressed in preimplantation mouse blastocysts and is present in the ICM (Niswander and Martin, 1992; Rappolee et al., 1994). *Fgf4* null mice are postimplantation-lethal due to impaired proliferation of the ICM (Feldman et al., 1995). *Fgf4* is also expressed in the primitive streak (E7.5-E8.5), and PSM and neuroectoderm in the trunk and continues to be expressed in the tail bud (Niswander and Martin, 1992). *Fgf8* transcripts are present in the epiblast and visceral endoderm at E5.75, and also localize to the primitive streak at E6.5 and subsequently in the tail bud (Crossley and Martin, 1995). *Fgf8* is found in the cells fated to become mesoderm and definitive endoderm (Crossley and Martin, 1995). *Fgf8* null mouse embryos show defects in mesoderm formation such that

epiblast cells accumulate in the posterior side of the embryo and fail to migrate away from the primitive streak through epithelial-mesenchymal transition (EMT) (Sun et al., 1999).

Consequently, *Fgf8*$^{-/-}$ embryos lose almost all tissues derived from mesoderm and definitive endoderm (Sun et al., 1999). The absence of *Fgf receptor 1 (Fgfr1)* in mouse embryos has a similar but less severe effect than seen in the *Fgf8*$^{-/-}$ embryos in that there is more development of mesoderm and endoderm tissues in *Fgfr1*$^{-/-}$ embryos (Partanen et al., 1998; Xu et al., 1999). Additionally, hypomorphic mutations in *Fgfr1* lead to posterior skeletal truncations of the embryos (Partanen et al., 1998). All of these analyses suggest Fgf signaling is critical to axial elongation of the mouse embryo after gastrulation.

The role of FGF signalling at later developmental stages during trunk and tail formation has been investigated through conditional mutants. *Fgfr1* conditional knockout induced by T-Cre recombinase in nascent mesoderm results in the truncations of the sacral and caudal regions of the vertebral column (Wahl et al., 2007). *Fgf4/Fgf8* conditional double knockout mutant embryos induced by the same T-Cre recombinase show severe axis truncation at the level of the forelimb buds, resulting from premature mesoderm differentiation (Naiche et al., 2011).

The *Fgf4/Fgf8* conditional double knockout mutant embryos appear superficially similar to the Wnt3a null mutants as described above, including a shortened axis and downregulation of T/Brachyury expression (Ciruna and Rossant, 2001; Takada et al., 1994; Yamaguchi et al., 1999). In addition, Wnt signaling is required for Fgf signaling expression during gastrulation. Fgf8 expression is down-regulated in β-catenin mutants at E6.5 (Morkel et al., 2003) and absent

from E8.5 embryos lacking β-catenin expression in the primitive streak using *T-Cre* (Dunty et al., 2008). On the other hand, *Wnt3a* expression is almost completely absent in *Fgf4/Fgf8* double mutants, which is likely due to the loss of posterior epiblast and nascent mesoderm in the mutants (Nowotschin et al., 2012). Overall, the phenotypic similarities of mutants in Wnt and Fgf signaling and the regulatory relationship suggest that these two pathways work together in promoting axial elongation and maintaining the active stem cell population in the epiblast (Boulet and Capecchi, 2012).

Retinoic acid signaling

Retinoic acid (RA) is the active derivative of vitamin A and is involved in several developmental processes, including axial elongation (Abu-Abed et al., 2001; Sakai et al., 2001). The biological function of RA relies on its two-step synthesis from either alcohol dehydrogenases (ADHs) or retinol dehydrogenase (RDHs) to produce retinaldehyde, and retinaldehyde dehydrogenases (RALDHs) to generate RA (Rhinn and Dollé, 2012). RA signaling regulates many developmental processes, including neurogenesis, cardiogenesis, forelimb bud initiation, eye development, vertebral patterning and body axis elongation (Duester, 2008). RA functions as a ligand for the nuclear RA receptors (RARα, RARβ, and RARγ), which form heterodimers with retinoid X receptors (RXRα, RXRβ, and RXRγ) to trigger the transcription of target genes (Rhinn and Dollé, 2012).

RA signalling is regulated by both RA synthesizing and catabolizing enzymes in a time- and tissue-specific manner. At early somite stages, RA is generated by RALDH2 in the somites and

anterior PSM and diffuses into the neural plate to ensure proper spinal cord neuronal differentiation (del Corral et al., 2003) and somitogenesis (Sirbu and Duester, 2006; Vermot et al., 2005) during body axis elongation. The oxidation of RA, which leads to its degradation, is accomplished by cytochrome P450 (CYP) proteins Cyp26A1, B1 and C1 (Ghyselinck and Duester, 2019). RA activity becomes progressively excluded from the posterior embryo with the onset of *Cyp26A1* expression in the tail bud (Abu-Abed et al., 2001; Rhinn and Dollé, 2012), thus the RA distribution is opposite to the *Fgf8* gradient during trunk formation (Fig. 2). Cells leaving the posterior RA-low progenitor zone in the tail bud are therefore exposed to higher RA and lower *Fgf8* levels (Figure 2), triggering their differentiation into the neural plate or segmentation into somites (del Corral et al., 2003). Exposure of E8.5 embryos to excess level of RA or inactivation of Cyp26A1 leads to severe axial truncations (Abu-Abed et al., 2001; Kessel and Gruss, 1991), while embryos lacking *Raldh2* exhibit axial truncation and die during development due to heart failure (Niederreither et al., 1999). These data show that RA distribution in the embryos is tightly regulated to ensure proper development.

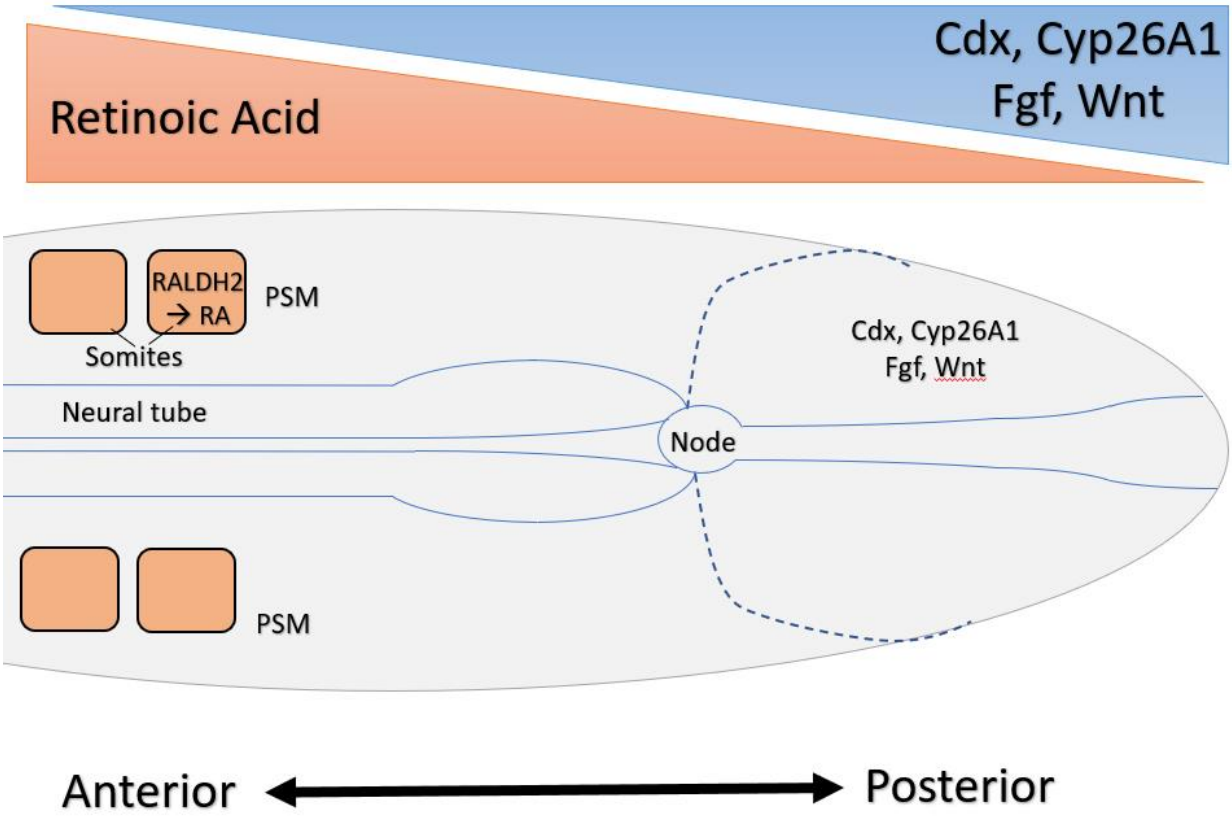


Figure 2. Schematic representation of Retinoic Acid, fibroblast growth factor (FGF), Wnt and Cdx gradients in vertebrate embryos undergoing somitogenesis.

T/Brachyury

At around E6.0 transcription factor *T/Brachyury* expression is initiated by Wnt3 in the primitive streak when the gastrulation and mesoderm formation initiates (Herrmann and Kispert, 1994). Later in development, *T/Brachyury* is expressed in the node and notochord and in the tail bud throughout axial elongation (Inman and Downs, 2006). The *T/Brachyury* mutant mice have been studied from the early 1900s. T/Brachyury heterozygous embryos exhibit a shorter tail, whereas the homozygous mutant embryos fail to grow an allantois and fuse with the chorion, resulting in embryonic death between E9.5 and E10.5 and exhibit a general defect in mesoderm production (Stott et al., 1993). In addition to this, homozygous null mutants show ectopic neural tube development and lack somites posterior to the seventh pair (Stott et al., 1993). T/Brachyury is also important for creating and maintaining the NMP cell niche through the direct regulation of expression of *wnt* ligands and of *cyp26a1* in zebrafish (Martin and Kimelman, 2010), and modulating the fate decision of NMP cells at the NSB and later in the CNH (Gentsch et al., 2013).

T/Brachyury is co-expressed with *Wnt3a* in the primitive streak and *T/Brachyury* null embryos present a similar phenotype to *Wnt3a* null mutants with severe axis truncation and ectopic neural structures, although *Wnt3a* null mutants don't exhibit defects in allantois development (Takada et al., 1994). The deficiency of paraxial mesoderm cells in *Wnt3a* mutants is consistent with the absence of *T/Brachyury* expression in these mutants (Yamaguchi et al., 1999). Expression studies have also shown that T/Brachyury directly interacts with the *Wnt3a* and *Fgf8* proximal promoters and their expression is downregulated in *T/Brachyury* mutant embryos (Evans et al., 2012). All these data suggest interactions between T/Brachyury, Fgf and Wnt signaling

pathways, as well as the clearance of RA in the posterior embryo by the T-mediated expression of *Cyp26A1* (Martin and Kimelman, 2010), are essential for the maintenance of NMP cells and the subsequent specification of their progeny necessary for axis elongation.

The Cdx Family

The homeobox gene *caudal* (*cad*) was originally identified in *Drosophila* and encodes a 427-amino acid homeodomain-containing transcription factor that binds DNA through an evolutionarily conserved α-helix domain (Levine et al., 1985). Homologues of Cad were subsequently identified in zebrafish (cdx1a, cdx1b and cdx4), mice (Cdx1, Cdx2, and Cdx4), and humans (CDX1, CDX2, and CDX4), among other species (Houle et al., 2003). *Cad* gene products are essential for evolutionarily conserved roles in the development of posterior embryos, and the loss of *Cad/Cdx* function typically results in anterior-posterior patterning defects and posterior truncation (Macdonald and Struhl, 1986; van Rooijen et al., 2012).

The murine Cdx family members exhibit overlapping and dynamic expression patterns (Fig. 3). *Cdx1* expression begins around E7.5 in the ectoderm and nascent mesoderm in a broad region of primitive streak (Meyer and Gruss, 1993). At the onset of neurulation (E8.5), *Cdx1* is expressed in a gradient from anterior low to posterior high with an anterior limit at the level of hindbrain and spinal cord boundary, while in the mesoderm *Cdx1* is expressed in somites, the proximal developing limb bud from E9.5 to E10.5 and the nephric cord and mesonephric ducts (Meyer and Gruss, 1993). Throughout these stages, *Cdx1* transcripts continue in the primitive streak region and subsequently in the tail bud, with eventual extinction of expression around E12. *Cdx1* re-

expresses in the hindgut endoderm at E12.5 and retains this expression postnatally (Meyer and Gruss, 1993).

Cdx2 expression initiates at E3.5 in the trophectoderm and is the only *Cdx* member expressed in extraembryonic ectoderm (ExE) during preimplantation development (Beck et al., 1995). *Cdx2* is required for differentiation of the trophoblast lineage essential for blastocyst implantation into the uterine wall (Chawengsaksophak et al., 2004; Chawengsaksophak et al., 1997; Strumpf et al., 2005). *Cdx2* expression in the embryo proper begins in the primitive streak at E6.5 with a posterior high to anterior low gradient (Beck et al., 1995; Foley et al., 2019). *Cdx2* is also expressed in the hindgut and cloacal endoderm, posterior neural plate, notochord and continues in the tail bud and eventually extinguished in this tissue at E12.5 (Beck et al., 1995; Foley et al., 2019). From this stage onward, *Cdx2* is expressed in the hindgut endoderm and subsequently the intestinal epithelium where it persists throughout adulthood (Beck, 2002).

Compared to *Cdx1* and *Cdx2*, *Cdx4* has a more restricted expression domain beginning at E7.5 in the allantois and the posterior primitive streak. At E8.5, *Cdx4* is expressed in the posterior neuroectoderm, PSM, lateral plate mesoderm and hindgut endoderm before extinction at around E13.5 (Gamer and Wright, 1993). *Cdx4* also exhibits a gradient expression in low anterior to high posterior with an anterior limit to the most recently formed somites (Gamer and Wright, 1993).

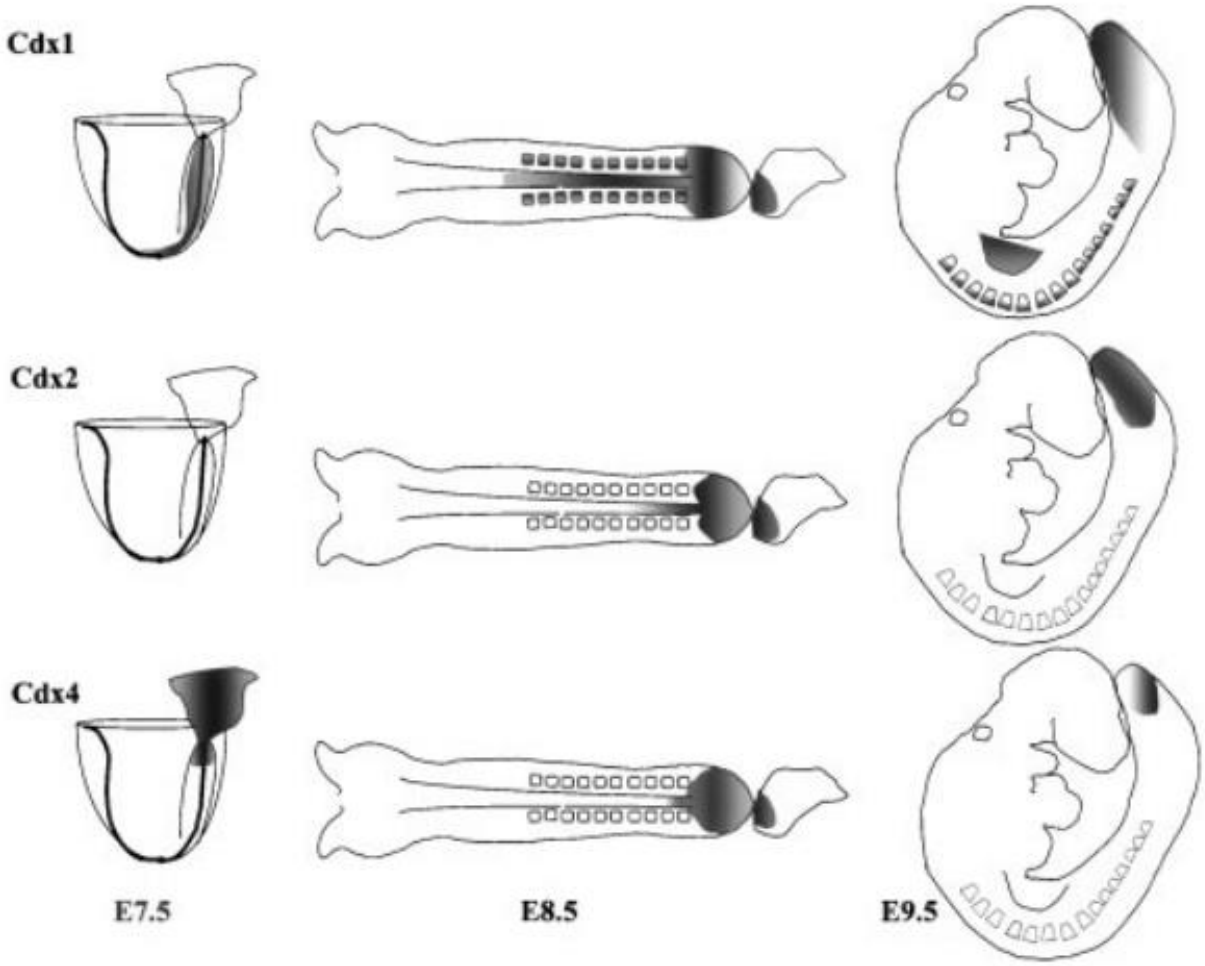


Figure 3. Representation of *Cdx* family members expression profiles in E7.5-E9.5 murine embryo. Cdx members are expressed in the embryos at different stages and intensities, which is indicated by the grey shading for each gene (Lohnes, 2003). (Reproduced with permission from the BioEssays publishing group).

Cdx loss-of-function studies

The role of Cdx family members in the mouse has been extensively investigated using loss-of-function models. *Cdx1*$^{-/-}$ mutants are viable and fertile but exhibit anterior homeotic transformations of the cervical and upper thoracic vertebrae related to altered somitic *Hox* gene expression (Subramanian et al., 1995). *Cdx1* heterozygous mutants also exhibit a subset of these defects with lower penetrance. Other tissues expressing *Cdx1* such as the limb buds, the intestine and the mesonephros appear unaffected (van den Akker et al., 2002). These observations show that *Cdx1* plays an important role in vertebral patterning *via* regulation of *Hox* gene expression.

Cdx2$^{-/-}$ mutants are peri-implantation lethal due to a critical role for Cdx2 in the trophectoderm. *Cdx2* conditional null embryos at E8.5 display caudal truncation, fail to develop hindlimb buds at later stages and die around E11.5 (Savory et al., 2009a). *Cdx2* heterozygotes are viable and fertile but have a shortened or kinky tail and anterior homeoses of the posterior cervical and anterior thoracic vertebrae reminiscent of the vertebral defects seen in *Cdx1*$^{-/-}$ mutants but extending to more posterior levels (Chawengsaksophak et al., 1997). These different axial levels of homeoses reflect the relative *Cdx1* and *Cdx2* expression patterns during development, with *Cdx1* expressed earlier and at a more anterior limit in the paraxial mesoderm and somites relative to *Cdx2*. These data imply that Cdx2 is essential for proper AP patterning and axial elongation.

In contrast to *Cdx1* and *Cdx2* mutants, *Cdx4* null mutant mice are phenotypically normal (van Nes et al., 2006). However, the loss of *Cdx4* exacerbates the phenotypes of both *Cdx1* and *Cdx2*

mutants, suggesting functional overlap among Cdx members (van Nes et al., 2006). Consistent with this, *Cdx1*$^{+/-}$*Cdx2*$^{+/-}$ and *Cdx1*$^{-/-}$*Cdx2*$^{+/-}$ compound mutants display more progressive vertebral defects and tail truncation compared to their single mutant counterparts (van den Akker et al., 2002). *Cdx2*$^{-/-}$*Cdx4*$^{-/0}$ double null mutants exhibit truncation of embryonic structures posterior to the hindlimbs, impaired chorio-allantoic fusion and subsequent placental disruption, resulting in the growth retardation and embryonic lethality at around E10.5 (van Nes et al., 2006). *Cdx1-Cdx2* conditional double mutants exhibit an open neural tube (craniorachischisis) and irregular and fused somites which are not found in neither single mutant (Savory et al., 2009a). Defects in the yolk sacs including disruption of blood formation and endothelial development are also evident in *Cdx1-Cdx2* double mutant embryos from E8.5 onwards (Savory et al., 2011b). Functional overlap among Cdx members is further supported by gene-swap studies, where *Cdx2* was knocked into the *Cdx1,* revealing functional equivalence between these paralogs as regards vertebral patterning (Savory et al., 2009b).

Cdx regulation of development through Hox genes

It is now established that the vertebral patterning role of Cdx is through the regulation of cohorts of *Hox* genes (Davidson et al., 2003; van Nes et al., 2006). Like *Cdx* members, *Hox* genes encode homeodomain-containing transcription factors that regulate AP patterning (Guazzi et al., 1994; Iimura and Pourquié, 2007; Wellik, 2007). In the mouse genome there are 39 *Hox* genes distributed in four chromosomal clusters: *Hoxa-Hoxd* (Fig. 4). Based on their homeodomain sequence and chromosomal localization within a cluster, *Hox* genes are further arranged in 13 paralog groups (Krumlauf, 1994). Murine *Hox* genes expression starts at E7.25 in the posterior

primitive streak and their expression is characterised by the correlation between physical chromosomal organization and spatio-temporal transcript distribution termed collinearity (Duboule, 1998; Duboule and Dollé, 1989; Krumlauf, 1994). This means that *Hox* genes located 3' within a cluster are expressed earlier (temporal colinearity) and with a more anterior boundary of expression (spatial colinearity) relative to progressively 5' *Hox* genes. *Hox* gene expression is observed in tissues from the hindbrain to the tail bud, including paraxial and lateral plate mesoderm, spinal cord, neural crest, limbs, hindbrain, surface ectoderm, branchial arches, and gut and gonadal tissues. This pattern of expression, together with numerous gain- and loss-of-function experiments, suggested an essential role for Hox proteins in specifying vertebral identities through a molecular "*Hox* code" comprised of distinct combinations of *Hox* gene products at a specific axial level at a given time (Kessel and Gruss, 1991).

A number of studies have shown that Cdx directly regulates *Hox* expression and several of the vertebral defects seen in *Cdx* mutants are close phenocopies to those of certain *Hox* mutants (Lohnes, 2003; Subramanian et al., 1995). Cdx transcription factors modulate gene expression by binding to *cis*-acting Cdx response elements (CDRE) at targets loci (Dearolf et al., 1989). A consensus CDRE has been defined of the sequence TTTATG, and this motif can be found in targets genes such as certain *Hox* loci, including the promoter regions for *Hoxa7*, *Hoxb8* and *Hoxc8* (Charité et al., 1998; Subramanian et al., 1995; Taylor et al., 1997). For example, the *Hoxa7* promoter contains two consensus CDREs, and the deletion of one of these elements attenuates response to CDX1 in transient transfection assays (Subramanian et al., 1995). In addition, the CDREs in the *Hoxb8* locus are necessary for the proper spatial expression of a transgenic reporter (Charité et al., 1998). These findings provide strong evidence of direct

regulation of certain *Hox* genes by Cdx transcription factors, and that this is the means by which Cdx members regulate vertebral patterning.

Hox genes have also been found to be implicated in the formation of primitive hematopoietic cells, yolk sac hematopoietic progenitors (Pineault et al., 2002). Studies in zebrafish utilizing morpholino-mediated (MO) knockdown revealed that compound *cdx1a*^{MO}*cdx4*$^{-/-}$ (analogous to murine *Cdx2*$^{-/-}$) mutants exhibit a complete lack of both primitive and definitive hematopoiesis, as well as perturbed expression of several *hox* genes (Davidson and Zon, 2006). Furthermore, ectopic expression of certain *hox* genes (e.g., *hoxa9a* and *hox7a*) can partially rescue blood formation, further supporting that *cdx* genes regulate hematopoiesis by the activation of downstream *hox* targets (Davidson et al., 2003; Davidson and Zon, 2006).

Murine embryonic stem cell (mESC) models have also been used to interrogate the role of Cdx in hematopoiesis. Cdx-deficient mESCs exhibited impaired embryonic hematopoietic progenitor formation, and a microarray analysis of *Cdx2*$^{-/-}$ mESCs revealed a significant reduction of *Hox* transcripts, due to Cdx2 deficiency (Wang et al., 2008).

Regarding the function of Cdx in axial elongation, the axial truncations seen in *Cdx2* mutants can be rescued by overexpression of certain trunk *Hox* genes (e.g., *Hoxa5* and *Hoxb8*). This complementation for *Cdx* deficiency indicates that *Hox* genes also participate in the axial elongation process as mediators of *Cdx* function (Young et al., 2009).

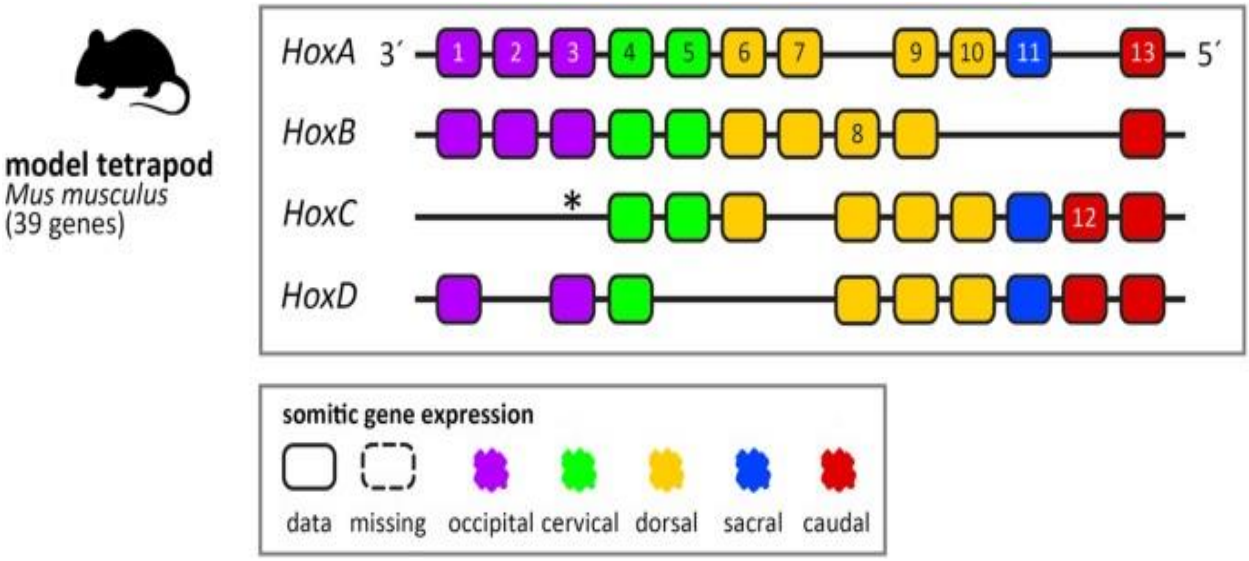


Figure 4. Overview of the murine *Hox* gene inventory with indication of somitic *Hox* gene expression. The 39 *Hox* genes in mice are arranged on four different chromosomes in four clusters (*HoxA*, *B*, *C* and *D*). The color coding indicates the *Hox* gene expression pattern in the axial regions (Böhmer and Werneburg, 2017). (Reproduced with permission from the Scientific Reports publishing group).

Development of a Cdx1/2 conditional mutant mouse model

As previously mentioned, *Cdx* mutant phenotypes clearly suggest a regulatory role for Cdx in AP vertebral patterning, but the functional overlap between Cdx members and the early lethality of the *Cdx2* null mutant impede further understanding of Cdx post-implantation function. To circumvent this limitation, our lab used the CRE/*loxP* recombinase system where the exon 2 of the *Cdx2* locus, which codes for most of the DNA binding homeodomain (Chase et al., 1999), was flanked by loxP sites and its excision is driven by the ubiquitously expressed *β-Actin* promoter. This transgene allows regulated activation of CRE through its fusion to an estrogen receptor ligand domain (ER^T) modified to respond exclusively to tamoxifen, with the fusion protein sequestered in the cytoplasm by Hsp90 in the absence of ligand binding. Tamoxifen treatment at post-implantation stages results in translocation of Cre-ER^T to the nucleus followed by Cre-mediated recombination and excision of *Cdx2* exon2 to inactivate the protein.

$Cdx2^{f/f}$ (i.e. *Cdx2* with exon 2 flanked by *loxP* sites) female mice were paired with $Cdx2^{f/f}$*; Cre-*ER^T males in timed matings and the pregnant female mice were dosed with 2mg of tamoxifen by oral gavage at E5.5. Analysis of mutants recovered from this approach revealed that Cdx2 impacts diverse developmental programs, including axial elongation, through non-*Hox* targets, such as *Cyp26a1* (Savory et al., 2009a). Thus, Cdx2 plays an important role in axial elongation through the direct regulation of both *Hox* and non-*Hox* targets responsible for the formation of paraxial mesoderm and axial elongation (Savory et al., 2009a).

Our lab has also generated *Cdx1/Cdx2* conditional double mutants to further address the issue of functional overlap between Cdx family members. The *Cdx2$^{f/f}$; Cre*-ER^{T} line was crossed into the Cdx1 null background (Subramanian et al., 1995), and pregnant females from *Cdx1$^{-/-}$Cdx2$^{f/f}$*;Cre-ER^{T} × *Cdx1$^{-/-}$Cdx2$^{f/f}$* matings were administered tamoxifen at E5.5 as above to inactive *Cdx2* post-implantation. The expression of *Cdx4* was found lost in *Cdx1/2* double knockout mutants (DKO) recovered at E8.5 (Savory et al., 2011b). These DKO mutants therefore represent a model of complete loss of Cdx function.

Cdx1/2 DKO embryos exhibit a more severe caudal truncation compared to *Cdx2* single null counterparts, as well as more progressive defects such as an open neural tube and irregularly shaped somites. Markers of mesoderm and paraxial mesoderm such as *T Brachyury, Wnt3a, Mesp1* and *Cyp26A1* were severely compromised in DKO mutant embryos at E8.5, as were players in planar cell polarity, consistent with the craniorachischisis phenotype (Savory et al., 2011b; Wallingford, 2006).

Rationale

Murine Cdx family members play crucial roles in vertebral AP patterning and axial elongation and are thought to exert their effects via direct regulation of direct targets, including *Hox* genes. While progress has been made, a better understanding of their function has been hampered by the peri-implantation lethality of *Cdx2* knockout and functional overlap between family members. Conditional deletion approaches allowed us to circumvent these limitations and lead to the identification of novel Cdx functions in development and non-*Hox* targets involved in

these processes (Savory et al., 2011a). My work focused on the application of a GFP transgenic mouse to enrich NMP cells, combined with RNA-seq and ChIP-seq approaches for the identification of novel Cdx target genes implicated in axial elongation.

Hypothesis

Murine Cdx members regulate axial elongation during embryo trunk and tail development through direct control of target genes which regulate NMP cell fate decision.

Objectives

Aim 1: Identify Cdx transcriptional candidate target genes that are implicated in axial elongation by RNA-seq.

Aim 2: Determine if the promoter of novel targets identified by RNA-seq are occupied by Cdx2 using ChIP-seq and ChIP-PCR.

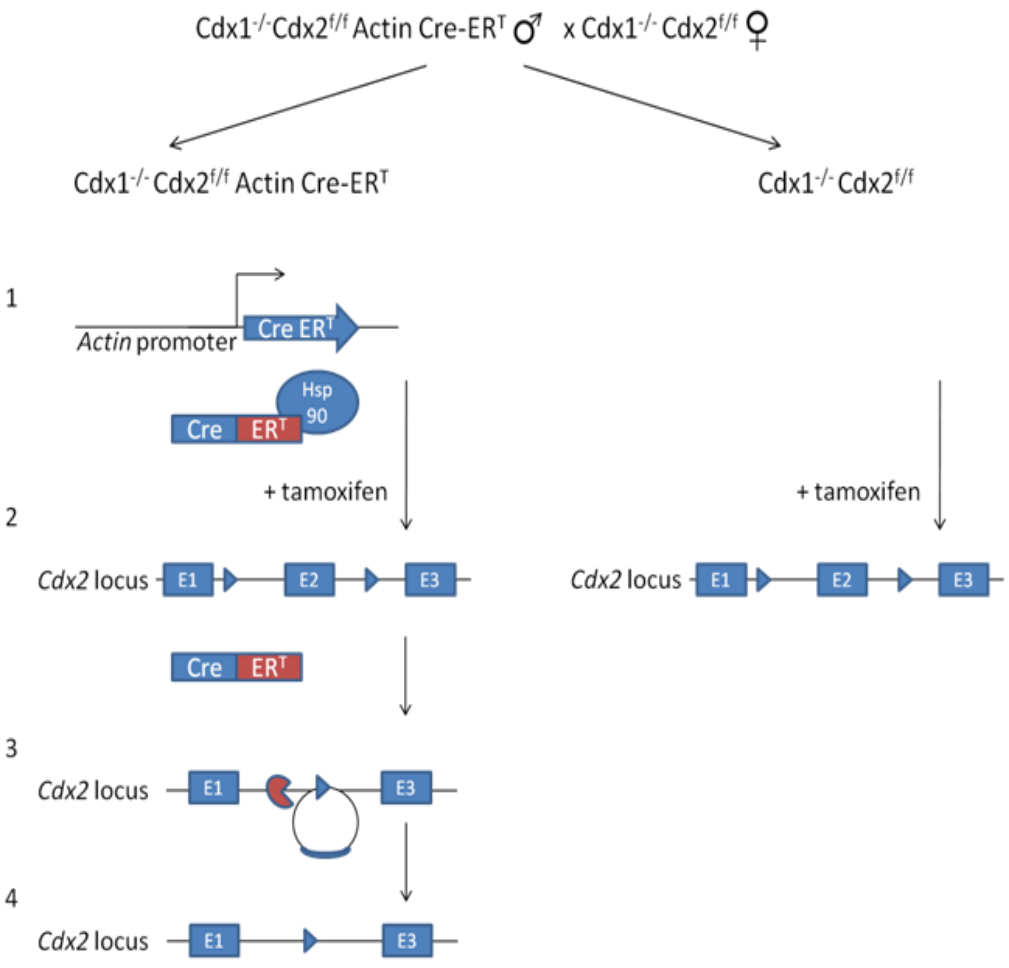
Aim3: Confirm the misregulation of the targets in Cdx null embryos by *in situ* hybridization.

Chapter 2

Materials and Methods

Mouse Model

To circumvent the pre-implantation lethality of the *Cdx2* null mutant and the overlapping function between *Cdx* family, I used our previously developed Cdx DKO conditional mutant model crossed with a *Pax2*- transgenic reporter expressing GFP in tail bud cells including NMP progenitors to better understand the role of Cdx in axial elongation (Figure 5).



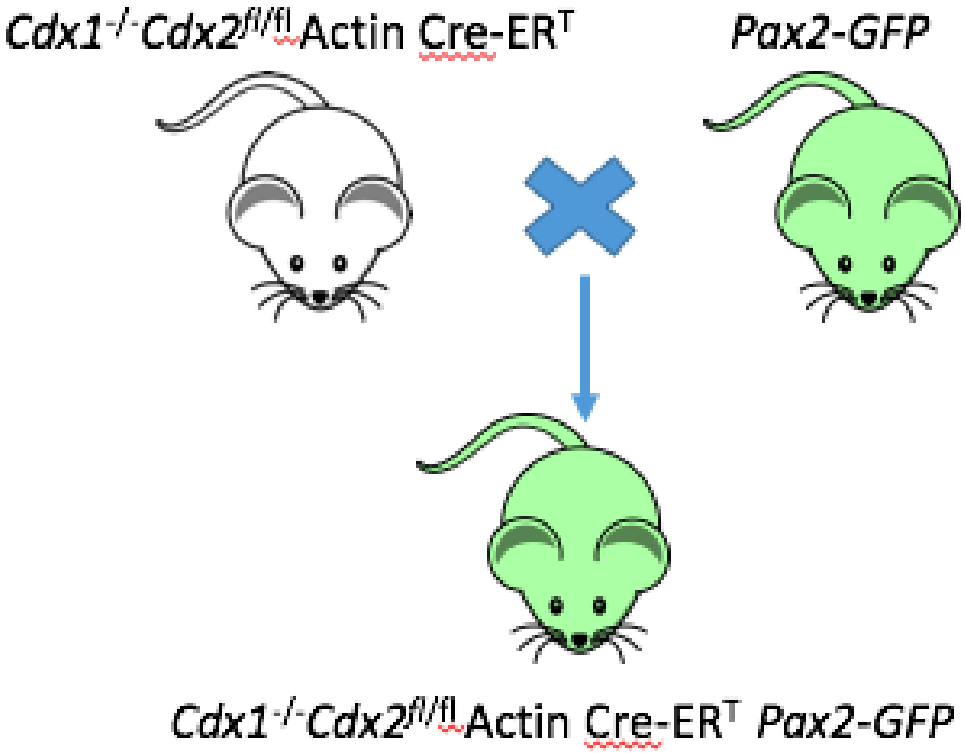


Figure 5. Mouse Model. Upon tamoxifen administration, the Cre-ER^{T} translocates to the nucleus and excises Exon 2 of Cdx2, which is flanked by LoxP sites, leading to a predicted non-functional protein. The Pax2-GFP model permits isolation of wild-type or Cdx-mutant GFP-positive cells anticipated to be enriched for NMPs and immediate progeny.

Embryo collection and analysis

CD-1 mice aged 6-8 weeks were obtained from Charles River. *Pax2*-GFP BAC transgenic mice and *Cdx1*$^{-/-}$*Cdx2*$^{f/f}$*Actin*Cre-ER^{T} animals have been previously described (Sharma et al., 2017; Sturgeon et al., 2011). Appropriate matings were used to generate *Pax2*-GFP; *Cdx1*$^{-/-}$*Cdx2*$^{f/f}$ *Actin*Cre-ER^{T} offspring. These were crossed with *Cdx1*$^{-/-}$*Cdx2*$^{f/f}$ females overnight, and noon on the day of detection of a vaginal plug was considered as E0.5. *Cdx2* deletion was effected by a 2mg dose of tamoxifen (Sigma Aldrich; dissolved in 5% ethanol in corn oil) at E6.5 or E7.5 by oral gavage. Embryos were harvested at E8.5-E9.5 for subsequent analysis using *Cdx1*$^{-/-}$*Cdx2*$^{f/f}$ littermates as controls unless otherwise noted.

Quantitative Real-Time Polymerase Chain Reaction (RT-qPCR) and semi-quantitative reverse-transcriptase PCR (RT-PCR)

RNA was isolated from embryos using Trizol (Thermo Fisher Scientific) as per recommendations, and complementary DNA (cDNA) generated using standard procedures. cDNA was subsequently amplified by PCR with GoTaq or qPCR with SYBR green (Promega) according to the manufacturer's instructions using oligonucleotides specific for *Evx1*, *Isl1*, *Sp8*, *Zic3*, *Nr2f1* or *Cyp26A1* with *β-actin* as an input control. qPCR was performed using a CFX96 Real Time PCR Detection System (Bio-Rad) and results analyzed by the 2-ΔΔCt method (Schefe et al., 2006) normalized to *β-actin*. For specificity, the dissociation curve was considered for each amplicon. For semi-quantitative analysis, PCR was performed over a series of cycles and

samples within the linear range of amplification used for analysis. Data in both cases is reflective of at least 3 independent biological samples.

Whole Mount in Situ Hybridization

Whole mount *in situ* hybridization was performed using *in vitro*-transcribed digoxigenin-labelled antisense RNA probes as previously described (Savory et al., 2009a). Briefly, E8.5 - E9.5 embryos were dissected out in ice cold PBS and fixed with 4% paraformaldehyde (Sigma) in PBS (PFA) at 4°C overnight. Embryos were washed in PBT (PBS containing 0.1% Tween-20), dehydrated with methanol and rehydrated with PBT. They were then treated with proteinase K (10 mg/ml in PBT) at room temperature, the reaction stopped with glycine (2 mg/ml in PBT) and embryos postfixed with 4% paraformaldehyde, 0.2% glutaraldehyde (Sigma). Hybridization was performed at 70°C overnight in hybridization buffer [50% formamide, 1.3 x SSC (3M NaCl, 300 mM Na citrate, pH 5.5), 5 mM EDTA, 0.2%Tween 20, 50 mg/ml yeast tRNA, 100 mg/ml heparin] containing the RNA probe, followed by three washes at 70°C in hybridization buffer. Embryos were then washed in TBST (25 mM Tris.HCl, pH 8.0, 140 mM NaCl, 2.7 mM KCl, 0.1% Tween 20), equilibrated with MABT (100 mM Maleic acid, 150 mM NaCl, 0.1 % Tween-20, pH 7.5), blocked with MABT containing 1% blocking reagent and 10% goat serum for 2-3h and incubated with a 1:2000 dilution of alkaline phosphatase-conjugated anti-digoxigenin antibody in MABT with blocking reagents at 4°C overnight. Embryos were washed extensively with MABT at room temperature for 24h, equilibrated in NTMT (100 mM Tris HCl, pH 9.5, 50 mM MgCl2, 100 mM NaCl, 0.1% Tween-20) and developed at room temperature with NBT/BCIP diluted in NTMT. Reactions were stopped with PBT, embryos fixed with 4% PFA and stored in PBT. Probes for *Evx1, Isl1, Sp8, Zic3* and *Nr2f1* were prepared from PCR-

amplified cDNA fragments cloned into appropriate vectors for *in vitro* transcription by using following primers (forward primer is the first of each pair):

Evx1 5′- ATGGAGAGCCGAAAGGACATGG-3′ and

3′- AGAGGAGGTGCCCCTCACCAGATAA-5′

Sp8 5′- ATGCTTGCTGCTACCTGTAATAAG-3′ and

3′- CACCGCAACGGCCTGGAGTGA-5′

Isl1 5′- CTCGAGATGGGAGACATGGGCGATCCA-3′ and

3′- GTAGCCAGTCCTATTGAGGCATGAGGATCC-5′

Zic3 5′- ATGACGATGCTCCTGGACGGA-3′ and

3′- CCAATTTTAACGAATGGTACGTCTGA-5′

Nr2f1 5′- CTCTCATCCGAGATATGTTGCTG-3′ and

3′- CGATTTGGAAGAGGACCATGAG-5′

Fluorescence-Activated Cell Sorting (FACS)

Tail buds from E9.5 mouse embryos were dissected in PBS and pooled in 10% bovine serum in PBS. Tail buds were then dissociated into a single cell suspension by trypsinization at 37°C for 15 min, centrifuged at room temperature for 5 minutes at 1200 rpm and washed twice with 10% bovine serum in PBS. FACS was performed on cells from both wild type and Cdx conditional

mutant tail buds based on GFP positivity, granularity and size (side scatter, forward scatter), and viability using 7-AAD. Sorted GFP+ cells were stored in RNA later (Thermo Fisher) until use.

RNA-sequencing

Pooled GFP$^+$ cells from E9.5 mouse tail buds were collected in 15ml Falcon tubes and pelleted by centrifugation, washed with cold PBS and resuspended in 100 ul of RNA later (Thermo Fisher) until RNA extraction. Total RNA was extracted from Pax2-GFP$^+$ cells sorted from approximately 15 pooled E9.5 control or *Cdx1*$^{-/-}$*Cdx2*$^{-/-}$ conditional mutant tail buds using the RNeasy mini kit (Qiagen), according to the manufacturer's recommendations and 500ng of total RNA was prepared for the RNA sequencing from each replicate. RNA integrity number (RIN) was determined with a Bioanalyzer (Agilent) and RNA-seq libraries were prepared and sequenced using HiSeq4000 by GenomeQuebec (Montréal, Canada). At least 25 M single end reads were generated for each library and aligned on the mouse mm10 genome assembly using TopHat 2.0.9 (Kim et al., 2013). TopHat output files were processed with SAMTools (Li et al., 2009) and BedTools (Quinlan, 2014) and coverages normalized per million reads mapped for each sample. For the replicates of pooled Pax2-GFP$^+$ cells RNA-seq samples, average coverage files were calculated from the normalized coverages of each replicate. DESeq2 was used to perform differential expression analyses with fragments per kilobase of transcript per million mapped reads log2-transformed after adding an offset of 1 to each value. The log2-transformed values were centered across samples before principal component analysis (PCA). The top 349 most variable genes that exhibited a log2Foldchange > 2 and q value < 0.05 in *Cdx1*$^{-/-}$*Cdx2*$^{-/-}$ conditional mutant compared to the controls were determined by row variance using the

genefilter::rowVars function and Gene Ontology (GO) analysis (Eden et al., 2007; Eden et al., 2009) used to identify enriched biological terms.

Chromatin immunoprecipitation (ChIP) analysis

ChIP assays from embryonic chromatin were performed as described previously (Savory et al., 2009b). Rabbit pre-immune serum (Santa Cruz) and ChIP against Wnt3a exon4 were employed as negative controls. PCR was directed over regions encompassing potential CDREs, or distal (control) intervals by standard methods using the following primers (forward primer is the first of each pair):

Evx1 5′-GCTGTACGGAGCACTTTCTG-3′ and 3′- ACTGACTGACGGTGTCCTCG-5′

Sp8 5′-GCGACATGGATGAGACAGAT-3′ and 3′- AACCAGACCAGCTGTATCTG-5′

Isl1 5′-GCAGAGAATCTGTTTCCCCG-3′ and 3′- GGAAACTGGTAGGAGATTTCG-5′

Zic3 5′-GAGTGGGAAGTTACTTGGTC-3′ and 3′- TCCAACTTCGTCGGAGTAGC-5′

Nr2f1 5′- CTATCCCGCCTAAAGCGTTAGC-3′ and 3′- AGGCAACAGCCTGTAAAGGC -5′

Image Processing and Analysis

All image processing, analysis and measurements were performed using Fiji (Fiji, RRID:SCR_002285). Data was analyzed and plotted using GraphPad Prism7 (Graphpad Prism, RRID:SCR_002798).

Chapter 3

Results

Cdx conditional mutants exhibit premature axial truncation

Elongation of the body axis is driven by progenitors residing in the primitive streak and later the NMP cells in the tail bud. A number of transcription factors and signaling pathways are known to be essential for axial elongation, including T/brachyury, and Wnt, Fgf and retinoic acid (RA) pathways (Boulet and Capecchi, 2012; Herrmann et al., 1990; Olivera-Martinez et al., 2012; Takada et al., 1994). Cdx function is also essential for axial elongation in part through regulation of expression of the RA catabolic enzyme Cyp26A1, thereby maintaining the RA-deplete environment necessary for the NMP niche (Martin and Kimelman, 2010; Savory et al., 2009a). However, our understanding of the nature of Cdx target genes involved in axial elongation is incomplete. To further address this, we used a conditional *Cdx1/2* conditional double null mutant line (Savory et al., 2011b). These mutants termed double knockout (DKO) for simplicity, also exhibit loss of *Cdx4* expression, which itself is a Cdx target gene (Savory et al., 2011a). This results in loss of all Cdx function, and DKO mutants exhibit axial truncation anterior to the forelimb bud at E9.5 (Savory et al., 2009a).

To isolate NMP progeny, *Cdx1*$^{-/-}$*Cdx2*$^{f/f}$*ActinCre*-ER^{T} mice were crossed with heterozygous mice harbouring one copy of a GFP reporter gene under the control of the *Pax2* promoter which labels lateral, and ventral mesoderm lineages as well as PSM and NMP cells in the tail bud (Sharma et

al., 2017). Pregnant females from this cross were treated with a single 2 mg dose of tamoxifen by oral gavage at E7.5 to generate DKO mutants, as previously described (Savory et al., 2011b); littermates without *ActinCre* were used as controls, as axis elongation is not compromised in *Cdx1*$^{-/-}$ embryos (Subramanian et al., 1995). GFP-labeled control or *Cdx* mutant embryos were harvested at E9.5, 48 hours after tamoxifen treatment. WT embryos showed GFP expression pattern primarily in the tail bud, as expected, while GFP positive cells were also observed in the presumptive residual tail bud in the mutant embryos (Fig. 6A-F).

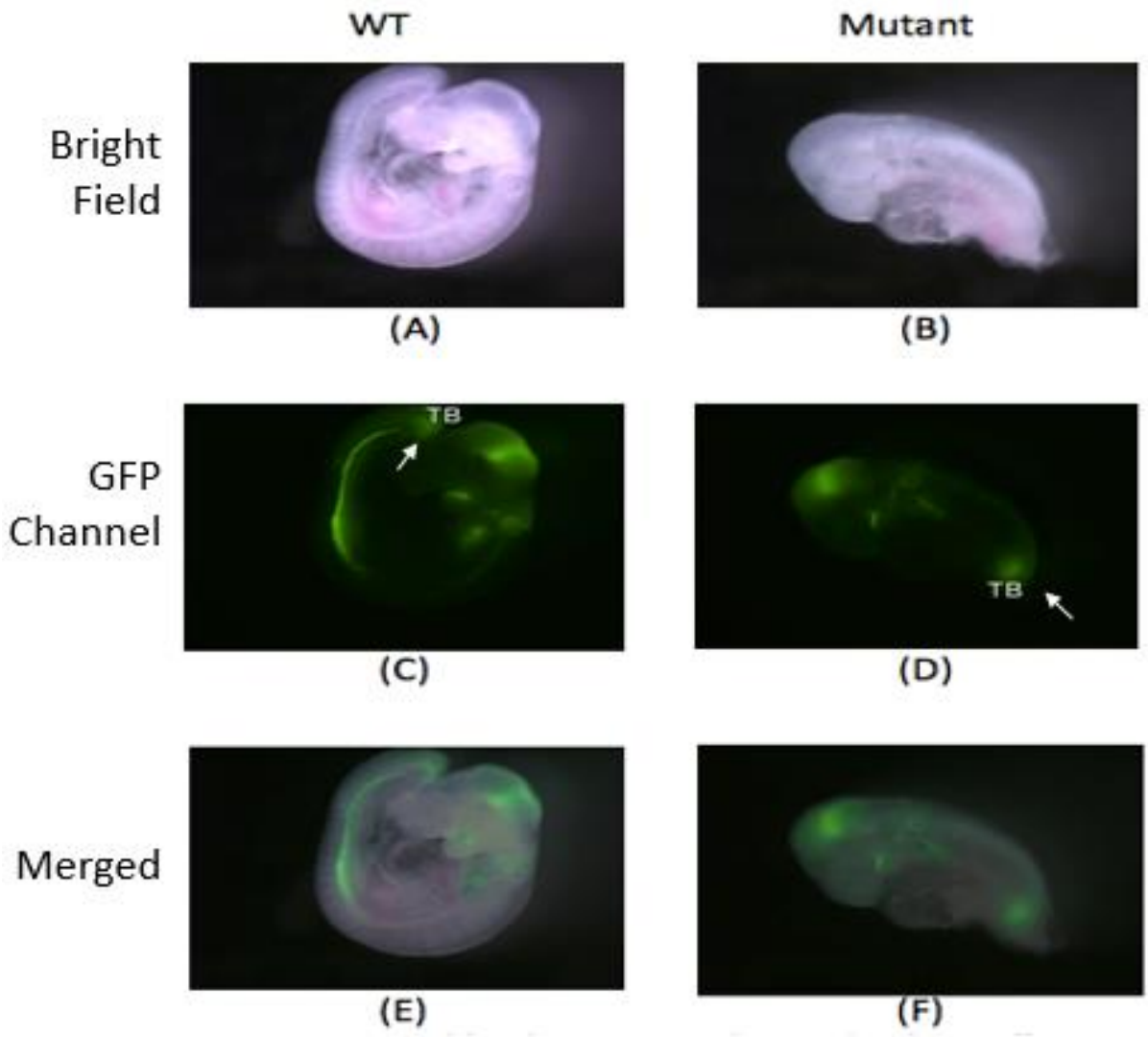


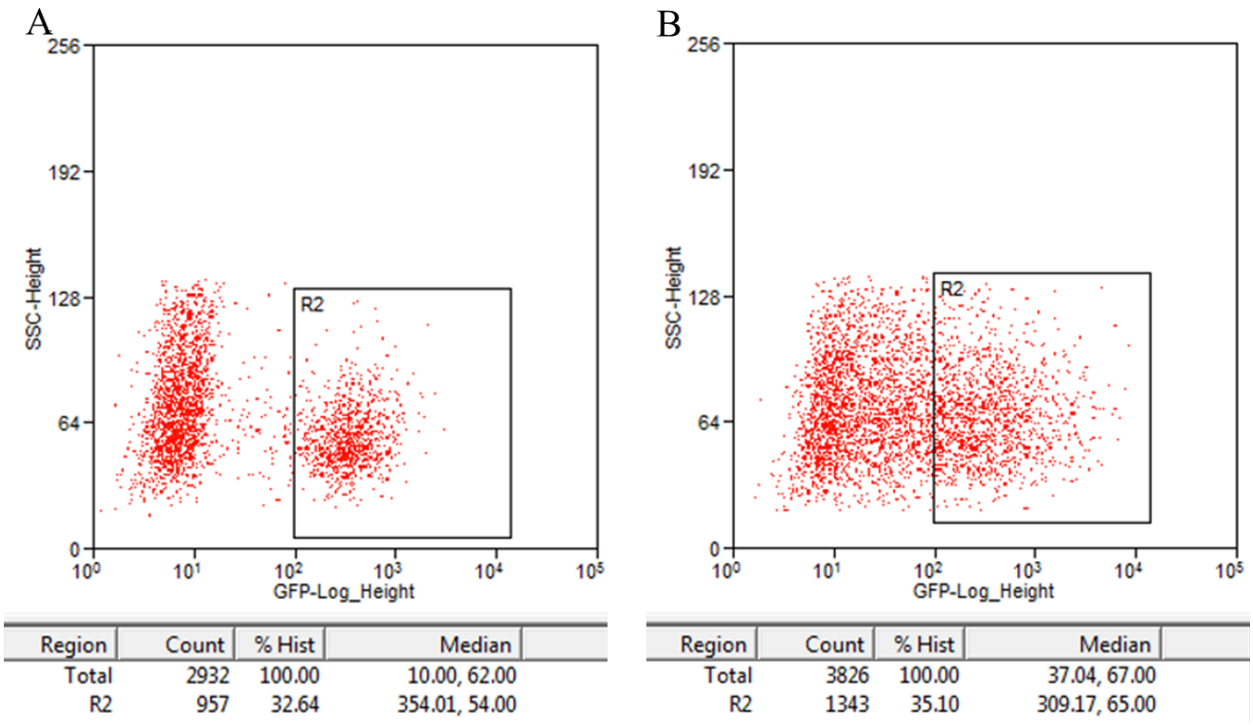
Figure 6. Cdx null mutants exhibit premature axial termination. (A-F) Pax2-GFP expression in E9.5 Cdx2 null mutant (B, D, F) and littermate *Cdx1*$^{-/-}$ control embryos (A, C, E). Pregnant females from *Cdx1*$^{-/-}$*Cdx2*$^{f/f}$*Cre*-ER^{T} *Pax2*-GFP x *Cdx1*$^{-/-}$*Cdx2*$^{f/f}$ matings were administered tamoxifen at E7.5 and embryos harvested 48H post-treatment. Fluorescence imaging shows GFP expression in the tail bud of both Cdx null (B, D) and littermate control (C, E) embryos (white arrows). TB, tail bud.

Identification of Cdx-dependent transcripts in the tail bud

RNA sequencing (RNA-seq) was performed on independent pools of GFP-positive cells from control and DKO tail buds to identify Cdx-dependent genes. GFP positive cells were enriched by fluorescence-activated cell sorting (FACS) from dissected tail buds of E9.5 control or *Cdx* mutant embryos. Notably, both control and *Cdx* mutant tail buds contained a similar percentage of GFP positive cells (Fig. 7A, B) although fluorescence was more variable in the cells from *Cdx* mutants, suggesting that this cell population may be losing GFP expression as axial elongation undergoes precocious termination.

RNA-seq from GPF^{+} isolates demonstrated reproducibility between *Cdx* null mutant and control samples by PCA, with a clear clustering along the principal component 1 (PC1) axis (83% variance) corresponding to their genotype (Fig. 8A). Fold-change of expression between *Cdx* mutants and controls was assessed using the MA-plot based method, with the fold change and the false discovery rate cutoffs set to 2 and 10%, respectively (Fig. 8B). The p-value histogram shows a peak at low p-values as expected for differentially expressed genes on a background of non-responsive transcripts with a flat distribution along the X axis. The horizontal green dashed line indicates the distribution of p-values under the null hypothesis of non-differentially expressed genes (Fig. 8C). The output of each mutant sample was normalized with the mean of the transcriptomes from control samples, revealing 1156 genes with altered expression in *Cdx* null mutant cells. These include 349 genes with an expression change greater than 2-fold, with a p-value cutoff set to 0.05. Among these were a number of genes involved in axial elongation and/or NMP differentiation including *T/brachyury*, *Wnt3a*, *Cyp26A1*, *Fgf8*, *Msgn1*, *Nkx1-2* and

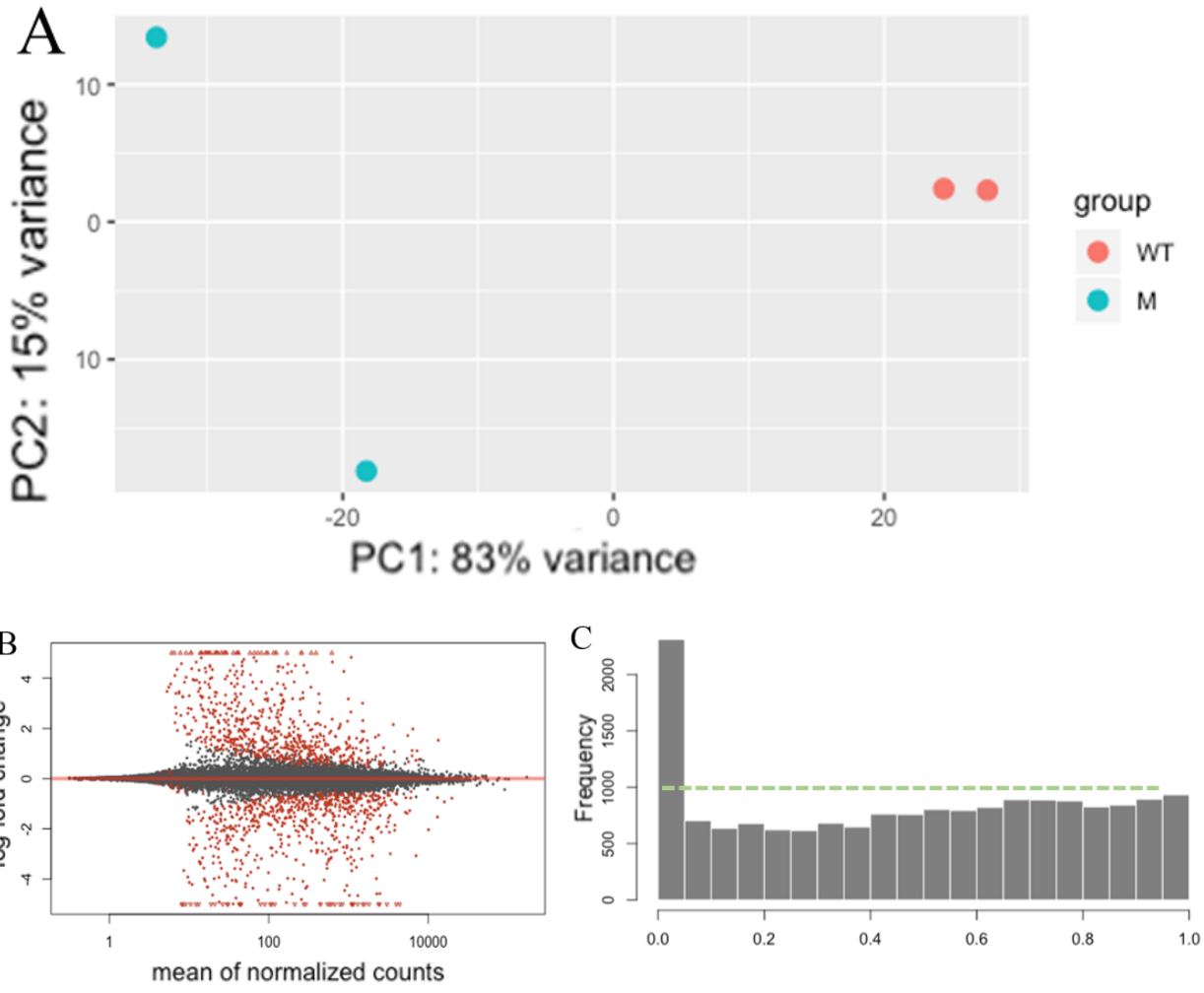
Tbx6, as well as certain posterior *Hox* genes (*Hox* paralog group *6-13*) (Fig. 8D). Gene ontology (GO) analysis of these 349 genes further revealed their involvement in pattern specification, regionalization, cell fate commitment and anteroposterior patterning (Fig. 9) consistent with the Pax2-GFP positive cells in the tail bud harboring a Cdx-dependent population that contributes to axial elongation.



Region	Count	% Hist	Median
Total	2932	100.00	10.00, 62.00
R2	957	32.64	354.01, 54.00

Region	Count	% Hist	Median
Total	3826	100.00	37.04, 67.00
R2	1343	35.10	309.17, 65.00

Figure 7. FACS sorting of GFP-expressing cells from tail bud tissue. Representative FACS plot of cells from control (A) and Cdx DKO mutant (B) tail buds from E9.5 embryos sorted based on GFP expression. The percentage of GFP-positive cells (rectangle) is 35.10% in Cdx null mutants and 32.64% in controls.


A
PC2: 15% variance
10
0
10
-20
0
20
PC1: 83% variance
group
WT
M
B
log fold change
4
2
0
-2
-4
1
100
10000
mean of normalized counts
C
Frequency
2000
1500
1000
500
0
0.0
0.2
0.4
0.6
0.8
1.0

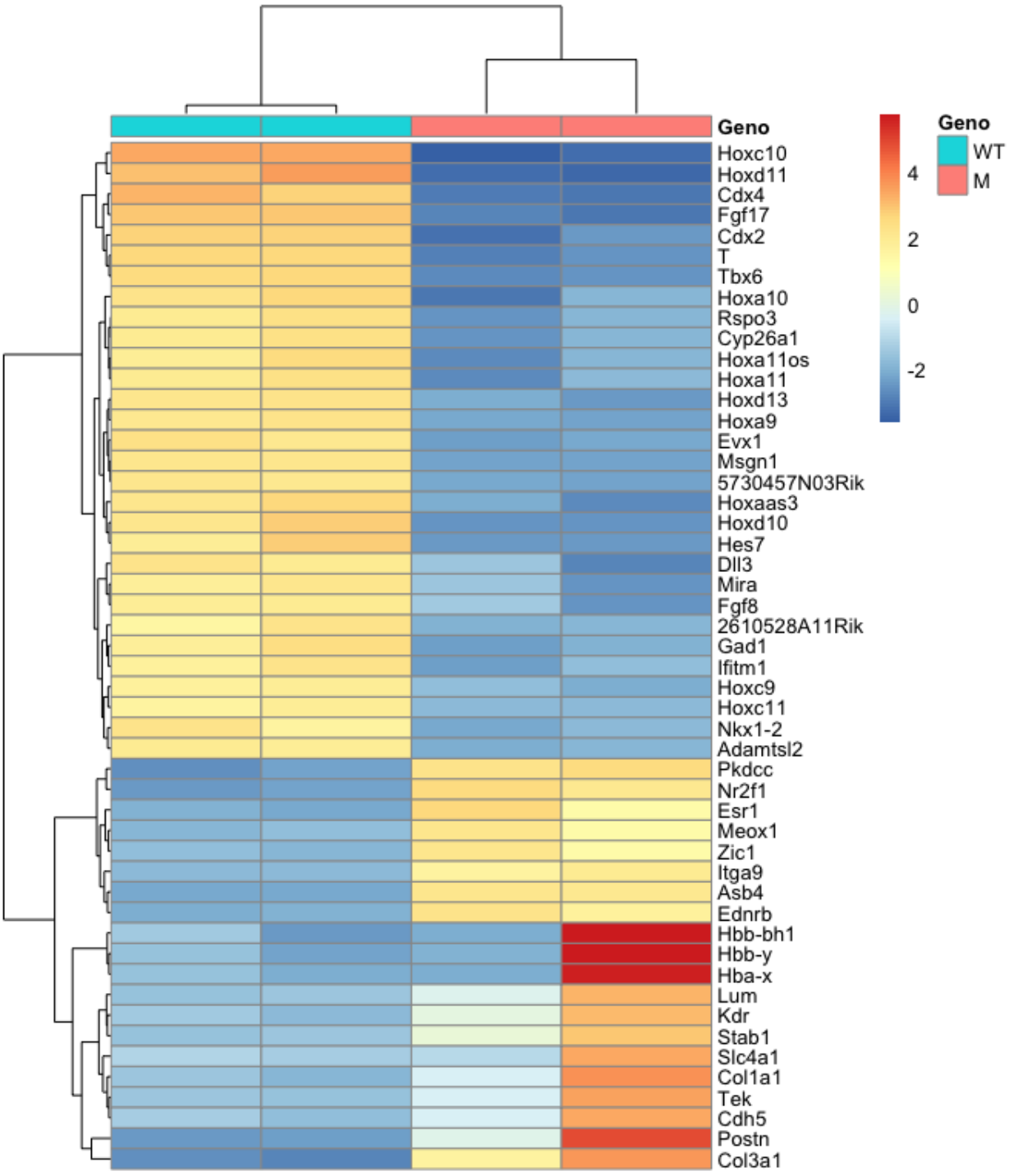

Geno
Hoxc10
Hoxd11
Cdx4
Fgf17
Cdx2
T
Tbx6
Hoxa10
Rspo3
Cyp26a1
Hoxa11os
Hoxa11
Hoxd13
Hoxa9
Evx1
Msgn1
5730457N03Rik
Hoxaas3
Hoxd10
Hes7
Dll3
Mira
Fgf8
2610528A11Rik
Gad1
Ifitm1
Hoxc9
Hoxc11
Nkx1-2
Adamtsl2
Pkdcc
Nr2f1
Esr1
Meox1
Zic1
Itga9
Asb4
Ednrb
Hbb-bh1
Hbb-y
Hba-x
Lum
Kdr
Stab1
Slc4a1
Col1a1
Tek
Cdh5
Postn
Col3a1
Geno
WT
M
4
2
0
-2

Figure 8. RNA-seq data analysis. (A) Principal component analysis (PCA) plot of RNA-seq datasets using pooled GFP^+ cells isolated from mouse E9.5 control (n=2) or Cdx mutant (n=2) tail buds. PC1 shows a strong difference between controls and Cdx mutant samples. (B) An MA-plot shows the fold change of gene expression between controls versus Cdx mutant cells with mean mRNA abundance (X axis). The red dots indicate genes with a greater than 2-fold alteration in expression. (C) P-value histogram for genes with mean normalized count greater than 1. (D) Heat map of scaled expression of the top 50 genes associated with axial elongation from controls and Cdx mutants.

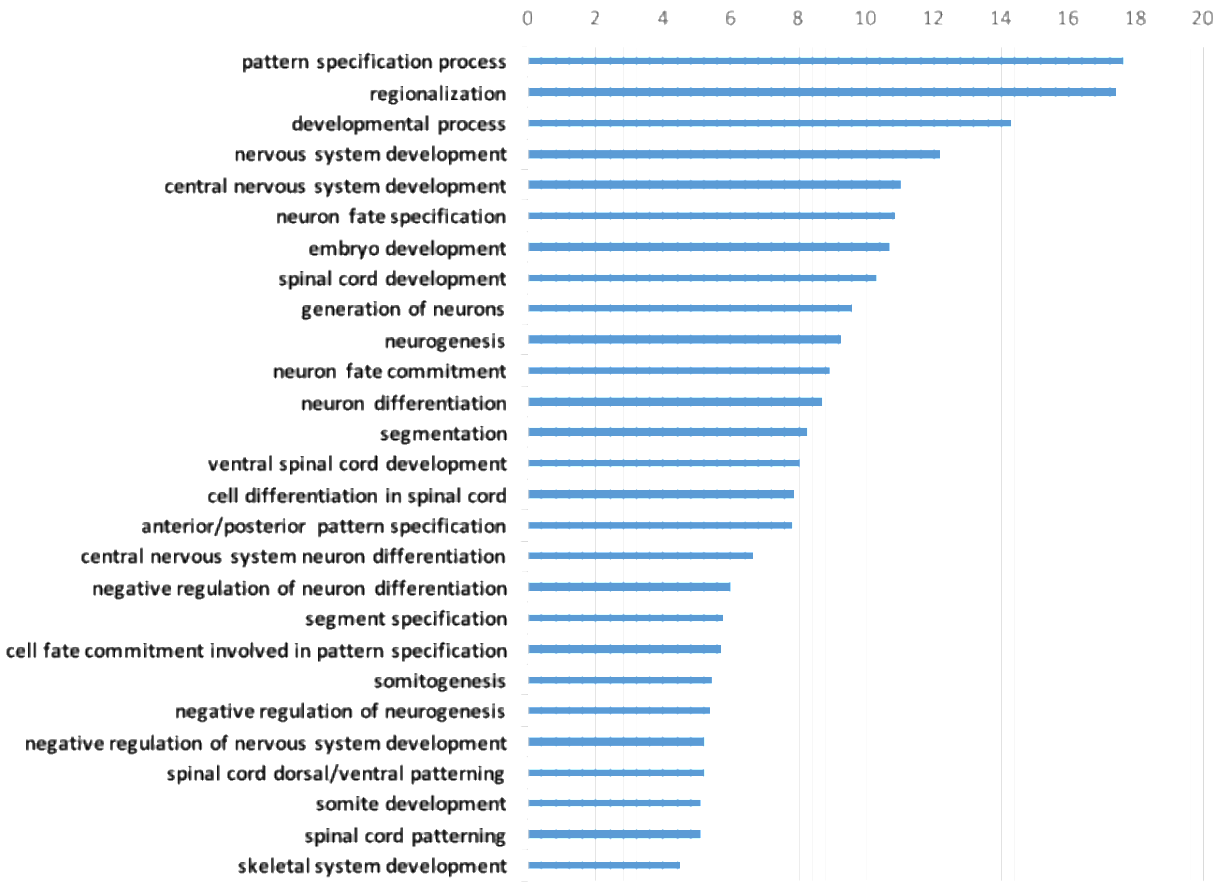


Figure 9. GO Enrichment Analysis. Bar charts showing the GO term enrichment analysis of the 349 genes for biological processes, ranked by fold enrichment and plotted as –log10(raw P-value).

Identification of novel Cdx targets involved in axial elongation

To identify potential direct Cdx targets, the Cdx-dependent genes in the RNA-seq dataset were cross-referenced to loci occupied by Cdx2 in E8.5 embryos as determined by chromatin immunoprecipitation-deep sequencing (ChIP-seq) (GEO accession GSE128858) (Foley et al., 2019). This yielded 33 Cdx-dependent transcripts with at least one Cdx2-bound region, 17 of which were downregulated and 16 upregulated in *Cdx* mutants compared to controls (Table 1). A number of these genes have previously been shown to be Cdx targets, including *Cyp26A1*, *Dll1*, *Lef1*, *Lmo2* (Brooke-Bisschop et al., 2017; Grainger et al., 2012; Mazzoni et al., 2013; Nguyen et al., 2017a; Savory et al., 2009a) thereby validating the fidelity of this exercise. A number of novel candidate target genes implicated in axial elongation were also recovered, including *Evx1, Sp8, Isl1* and *Zic3* (which were downregulated in the *Cdx* mutant tail buds) and *Nr2f1* (which was upregulated in the *Cdx* null mutants); all contained at least one CDRE associated with Cdx2 occupancy (Fig. 10A).

For further analysis of the *Cdx*-dependent regulation of *Evx1, Sp8, Isl1, Zic3* and *Nr2f1*, the Cdx2 binding peaks of these genes were assigned to high-throughput Chromosome Conformation Capture (Hi-C seq) data to attempt to reconcile the finding that some of the Cdx2 binding peaks are distal to the relevant transcription unit (e.g., the Cdx2 binding peak for *Zic3* is 120Kbp away from the transcription start site). It is possible that it is located in the same topologically associated domain (TAD), which can be viewed as an indicator of how likely two pieces of DNA are spatially proximal to each other. Currently one of the best characterized mechanisms for large domain organization and long-range chromatin interactions involve the

proteins CCCTC-binding factor (CTCF in mammals) and Condensin respectively. CTCF is a critical transcriptional regulator that is often enriched at the border of TADs (Vietri Rudan et al., 2015), whereas Condensin is a member of the structural maintenance of chromosomes (SMC) family that promotes the formation of TADs and chromosomal territories and mediate enhancer-promoter contacts (Iwasaki et al., 2019). The Hi-C seq analysis of relevant Cdx2 ChIP-seq peaks along with CTCF and Condensin II tracks from available data sets show that Cdx2 binding and the corresponding regulated genes are typically resident in the same TAD, and in some cases the binding sites of Cdx2 colocalize closely with CTCF and Condensin II (Fig. 11). While the Hi-C seq data was generated from ES cells (Dixon et al., 2015) and not E9.5 mouse tail buds, TADs are typically similar between different tissues of the same organisms (Dixon et al., 2012). In this regard, although *Evx1* is not considered to be co-regulated with the *HoxA* cluster, the Cdx2 binding region of *Evx1* and *HoxA* clusters are in the same 5' sub-TAD and exhibit strong interactions.

Table 1. Genes that are found in both RNA-seq and ChIP-seq.

Downregulated in Mutants	Upregulated in mutants
Evx1	Nr2f1
T	Zic1
Cdx2	Pou3f2
Sall3	Meis1
Cyp26a1	En2
Lef1	Pou3f3
Dusp4	Dbx1
Hoxd1	Ebf3
Mnx1	Irx5
Mllt3	Pax1
Lmo2	Tshz3
Zic3	Meis2
Isl1	Gfra1
Dll1	Sox21
Sp8	Ascl1
Sp5	Irx3
Arl4d	

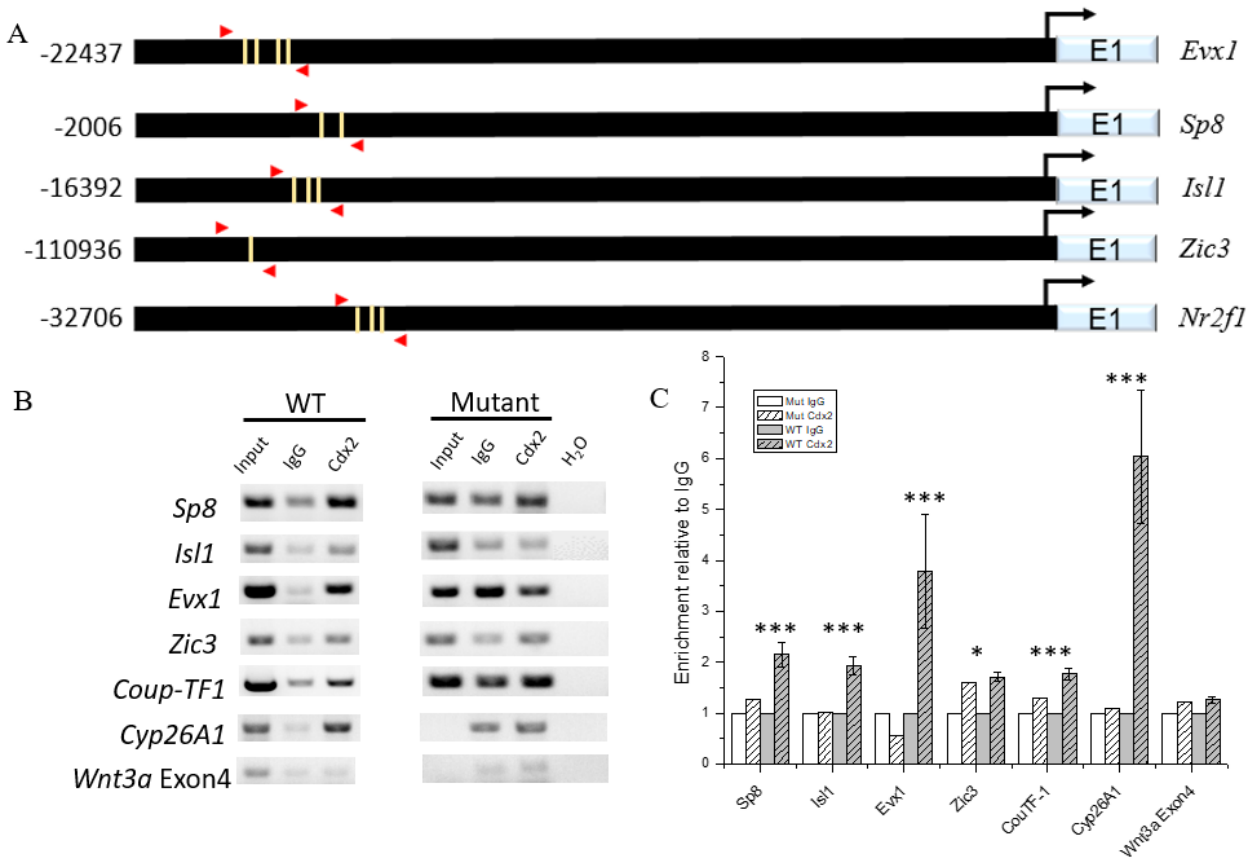


Figure 10. Enriched Cdx2 binding at the loci of candidate Cdx targets. (A) Schematic representation of CDREs in the promoter of candidate Cdx targets. Yellow lines indicate the position and sequence of the putative CDREs of 5 potential target genes. Red arrowheads are the positions of the prime used in ChIP assay. Black arrows are the transcription start sites. E is exon 1 of each gene. (B, C) Association (B) and quantification (C) of Cdx2 within the genomic loci of Sp8, Evx1, Isl1, Nr2f1 and Zic3 by ChIP-PCR using primers flanking peaks from ChIP-seq analysis. Error bars represent s.e.m. from three biological replicates. Cyp26A1 was used as a positive control and Wnt3a Exon4 as a negative control. ***P≤0.01, *P≤0.05 (determined by student test).

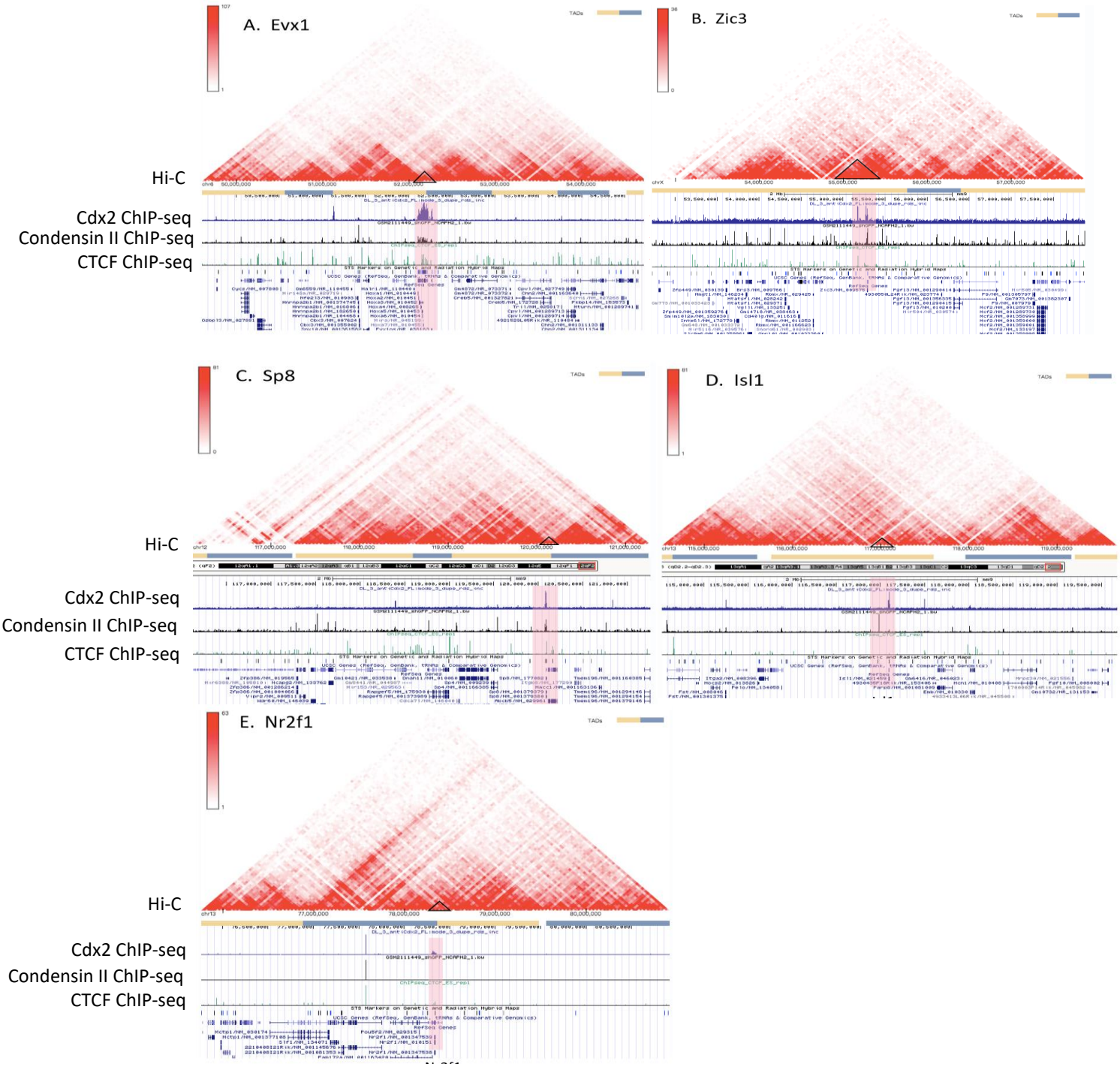

A. Evx1
B. Zic3
C. Sp8
D. Isl1
E. Nr2f1
Hi-C
Cdx2 ChIP-seq
Condensin II ChIP-seq
CTCF ChIP-seq
Hi-C
Cdx2 ChIP-seq
Condensin II ChIP-seq
CTCF ChIP-seq
Hi-C
Cdx2 ChIP-seq
Condensin II ChIP-seq
CTCF ChIP-seq

Figure 11. Normalized Hi-C heatmaps of chromatin interaction for candidate Cdx targets at 40kb resolution. Chromatin interactions in mESCs aligned with Cdx2 ChIP-seq in mouse 8.5 embryos, Condensin II ChIP-seq and CTCF ChIP-seq in mESCs. Black triangle indicates a sub-domain and pink rectangles represent the Cdx2 binding sites of the specific candidate targets.

Validation of Cdx2 targets

To further assess Cdx-dependent regulation of the candidate targets *Evx1, Sp8, Isl1, Zic3* and *Nr2f1*, their expression was assessed using in situ hybridization (ISH). Consistent with the RNA-seq analysis, expression of *Evx1, Sp8, Isl1* and *Zic3* was decreased, while *Nr2f1* was increased, in the caudal region of E9.5 Cdx mutant embryos compared to littermate controls (Fig. 12A). Cdx-dependent regulation of *Evx1, Sp8, Isl1, Zic3* and *Nr2f1* was also confirmed by RT-qPCR from E9.5 Cdx null mutant tail buds compared to controls (Fig. 12B). Similar regulation was also observed at E8.5, which corresponds to the period of trunk tissue generation (Fig. 12C), and prior to overt morphological defects in Cdx mutants, consistent with direct regulation of these genes by Cdx.

ChIP-PCR was conducted to validate Cdx2 enrichment within the genomic loci of the candidate target using E9.5 embryos and primers flanking putative CDRE identified by motif analysis under binding peaks recovered in ChIP-seq (Fig. 10B, C). This analysis confirmed Cdx2 binding to *Evx1, Sp8, Isl1, Zic3* and *Nr2f1* loci. These data collectively support a role for Cdx in directing expression of these genes during axial elongation.

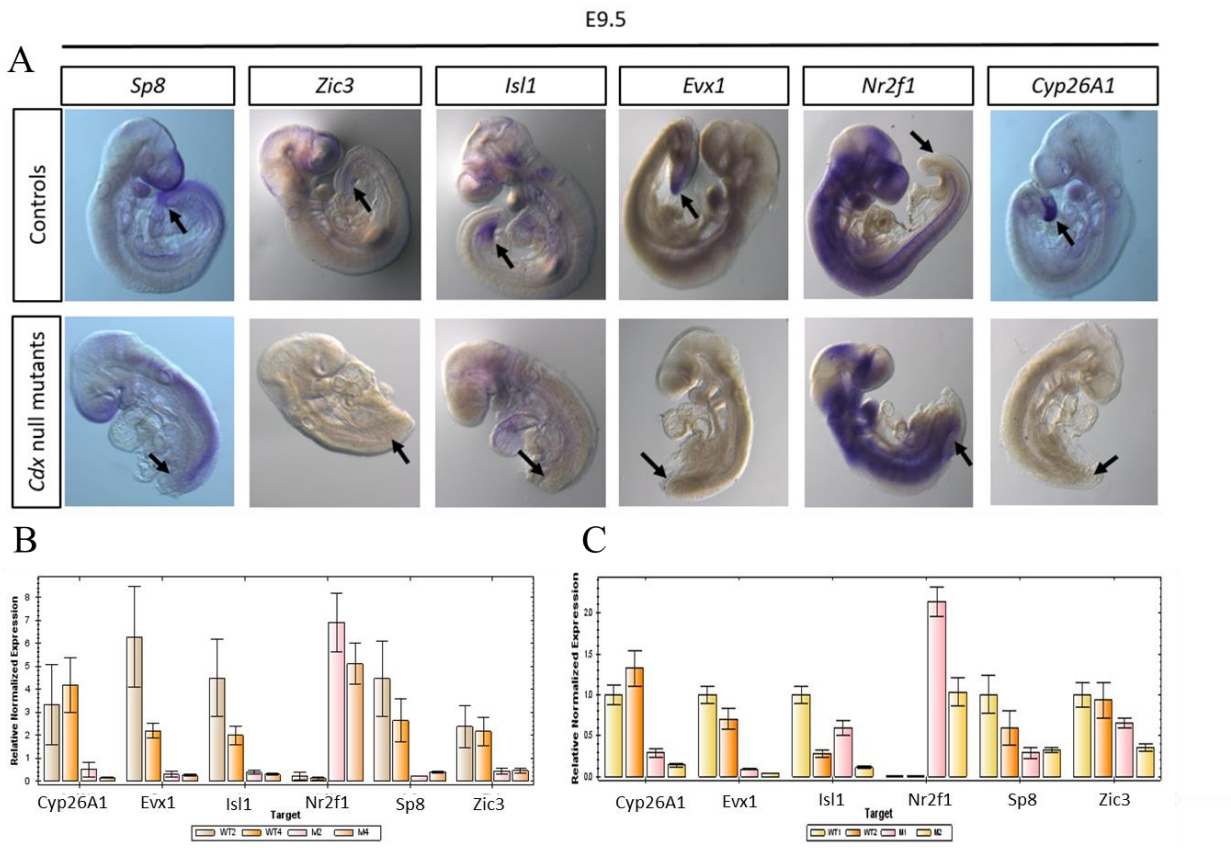


Figure 12. Characterization of candidate Cdx targets. (A) *In situ* hybridization analysis of Sp8, Evx1, Isl1, Nr2f1 and Zic3 in E9.5 control versus Cdx mutant embryos. Images are representative of 3 replicates for each genotype. Arrows indicate the tail buds in the controls or the residual tail bud region in the mutants. (B, C) RT-qPCR analysis for Sp8, Evx1, Isl1, Nr2f1 and Zic3 in tail buds from E9.5 (B) or E8.5 (C) control or Cdx mutant embryos.

Cdx targets are expressed within NMP cells

NMP cells are thought to arise at ~E7.5 in the node-streak border (NSB) and adjacent caudal lateral epiblast of the mouse embryo and are later located in the chordoneural hinge of the tail bud (Cambray and Wilson, 2007; Henrique et al., 2015; Tzouanacou et al., 2009). To address whether the newly identified target genes are co-expressed with *Cdx2* in NMPs, we analyzed publicly available single cell RNA-seq (scRNA-seq) from E7.0 embryos, E8.0 caudal lateral epiblast cells or E9.5 posterior embryos (Dias et al., 2020; Diaz-Cuadros et al., 2020; Scialdone et al., 2016) (Fig. 13). Although NMP cells do not express any singular unique marker, they can be characterized by the co-expression of *T/brachyury, Sox2*, and *Nkx1-2* (Rodrigo Albors et al., 2018). Using these criteria, we found that *Evx1*, *Sp8*, *Isl1* and *Zic3* are co-expressed with *Cdx2* in cells expressing the above NMP signature at all three time points with a Pearson coefficient > 0.93. Among these, *Zic3* is expressed with *Cdx2* in NMPs with a decreasing trend in E9.5 posterior embryos, consistent with a role for *Zic3* in promoting neural fate from NMPs (Henrique et al., 2015). Conversely, *Nr2f1* is not present in the NMP niche, consistent with the RNA-seq results, and suggesting Cdx represses more anterior character promoted by NR2F1 (Qiu et al., 1997) in the NMP population.

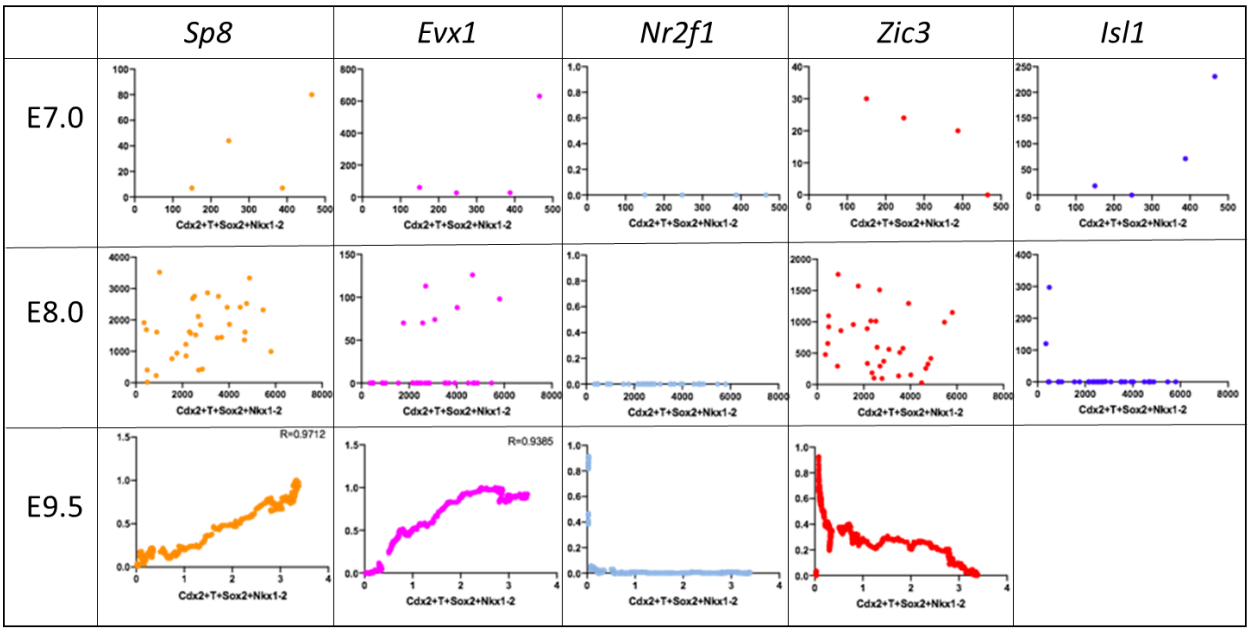


Figure 13. Candidate target genes are co-expressed with Cdx2 in putative NMPs. Single cell RNA-seq data from mouse E7.0 embryos, E8.0 caudal lateral epiblast cells or E9.5 posterior embryos (Dias et al., 2020; Diaz-Cuadros et al., 2020; Sanchez-Ferras et al., 2016) was used to assess the correlation between cells expressing Cdx2 and *T/brachyury* + *Sox2* + *Nkx1-2* with *Sp8*, *Evx1*, *Isl1*, *Nr2f1* and *Zic3*. R is the Pearson correlation coefficient.

Chapter 4

Discussion

Cdx family members are evolutionarily conserved regulators of AP patterning and caudal development during embryogenesis (Chawengsaksophak et al., 1997; van den Akker et al., 2002). Several independent studies have also shown that the Cdx proteins control axial elongation in a dosage-dependent manner (Savory et al., 2011b; van den Akker et al., 2002; van Nes et al., 2006). The identification of Cdx target genes is essential to better understand the functions of Cdx in developmental processes. Previous work in our lab suggested that *Cdx2* impacts axial elongation through the regulation of the expression of several genes, including *Wnt3a*, *T/brachyury* and *Cyp26A1* (Savory et al., 2009a), and there are likely a number of additional uncharacterized targets. However, the functional redundancy between Cdx members (Savory et al., 2011a; Savory et al., 2009b; van den Akker et al., 2002) and the peri-implantation lethality of homozygous *Cdx2* mutants (Chawengsaksophak et al., 1997) made it difficult to further study the function of the Cdx family. The generation of conditional *Cdx1/2* double mutants by our lab circumvented these limitations; this double knockout mouse line also exhibited loss of Cdx4, thus these mice recapitulate complete loss of Cdx function (Savory et al., 2011a). The additional use of a Pax2-GFP (Sharma et al., 2017) transgenic line allowed the enrichment of progenitor populations by FACS from tail bud cells. Using these isolates for RNA-seq combined with ChIP-seq data, qRT-PCR, in situ hybridization, ChIP-PCR and single-cell RNA-seq analyses suggest that Cdx impacts axial elongation through the regulation of target gene expression, including the novel targets identified in this work: *Evx1, Sp8, Zic3, Isl1* and *Nr2f1*. My studies expand on the proposed central role for Cdx in integrating multiple signaling

networks to orchestrate the maintenance of axial progenitors and their subsequent differentiation necessary for axial extension.

Identification of novel Cdx target genes implicated in axial elongation

RNA-seq using RNA extracted from Pax2-GFP fluorescence cells dissected from mouse tail buds revealed a set of genes that change their expression in Cdx1/2 double mutants, which could shed light on the role of Cdx in axial elongation. This analysis uncovered 349 differentially expressed genes with more than a 2-fold change following loss of Cdx. These genes encode products that are predominantly involved in AP patterning, regionalization, neuron fate specification and other developmental processes as revealed by GO analysis, suggesting that Cdx function coordinates gene regulatory networks that are essential for developmental processes. In addition, Cdx2 binding in the genome by ChIP-seq in E8.5 mouse embryos (Foley et al., 2019) identified 33 of the 349 genes that are occupied by Cdx2, some of which have previously been shown to be Cdx target genes (Amin et al., 2016; Brooke-Bisschop et al., 2017; Grainger et al., 2012; Savory et al., 2009a; Savory et al., 2011b) supporting the validity of the novel direct candidate targets of Cdx. 5 of these 33 potential direct target genes, *Sp8, Evx1, Zic3, Isl1* and *Nr2f1* were selected for additional analysis based on their known roles in processes related to axial elongation (Béland and Lohnes, 2005; Inoue et al., 2007; Jurberg et al., 2013; Kennedy et al., 2016; Prajapati et al., 2020).

According to RNA-seq, *Sp8, Evx1, Zic3* and *Isl1* were downregulated, and *Nr2f1* upregulated, in the Cdx null mutant tail buds. Whole mount in situ hybridization and RT-qPCR analyses showed that *Sp8, Evx1, Zic3* and *Isl1* are expressed in the WT caudal embryonic tail bud tissues and their expression is attenuated in *Cdx* mutants, while *Nr2f1* is absent in WT tail buds and upregulated in the residual posterior GFP+ cells in mutants, consistent with RNA-seq analysis. Cdx2 occupancy at the loci of *Sp8, Evx1, Zic3, Isl1* and *Nr2f1* was confirmed by chromatin immunoprecipitation (ChIP-qPCR) in WT mouse embryos, and this occupancy was lost in Cdx DKO mutant embryos. These results indicate that Cdx mediates certain of its effects on axial elongation through regulation of expression of the direct targets *Sp8, Evx1, Zic3, Isl1* and *Nr2f1*.

Cdx2 directly activates genes involved in axial elongation in NMPs

Axial elongation proceeds through the coordinated proliferation and subsequent differentiation of bipotential cell populations from the posterior growth zone located in the node/primitive streak, and later in the tail bud. This region comprising the caudal node and anterior most PS is called the node-streak border (NSB), and this bipotential NMP cell population processes the ability to self-renew and generate stem cells of both mesodermal and neural lineages, such as skeletal muscle and spinal cord derivatives. NMP cells are thought to arise at ~E7.5 in the NSB and adjacent caudal lateral epiblast (CLE) of the mouse embryo, and the derivatives of the NSB are later located in the chordoneural hinge (CNH) of the tail bud until axis elongation ceases around E13.5 (Cambray and Wilson, 2007; Henrique et al., 2015; Tzouanacou et al., 2009; Wymeersch et al., 2016). Throughout these developmental stages, NMPs are characteristically marked by the

co-expression of *T/brachyury*, *Sox2* and *Nkx1-2* (Rodrigo Albors et al., 2018) and integrate complex signaling and transcriptional events to maintain a balance between the undifferentiated progenitor state and differentiation into mesoderm and neuroectoderm.

Various components impact NMPs include Fgf, Wnt3 (Wnt/β-catenin), RA signaling and Cdx transcription factors (Chawengsaksophak et al., 2004; Savory et al., 2009a; Takemoto et al., 2006). Loss of both Fgf4 and Fgf8 in early gastrulation leads to embryonic lethality with mutants showing axial truncation which are consistent with disruptions in NMPs (Naiche et al., 2011). In addition, Fgf8 and Fgfr1 null mutants exhibit ectopic neural tubes in place of presomitic mesoderm, suggesting NMPs have shifted their fate from mesoderm to neural lineages in the absence of Fgf signalling (Boulet and Capecchi, 2012). Evidence also suggests that Wnt/β-catenin is important for maintenance of NMP cells. For example, the number of NMP cells was remarkably reduced when the level of β-catenin was attenuated during axial elongation (Wymeersch et al., 2016). As previously discussed, Cdx null mutants exhibit posterior elongation defects and attenuation of genes encoding Fgf and canonical Wnts in the growth zone in (Chawengsaksophak et al., 2004; Savory et al., 2009a; Savory et al., 2011b). Cdx appears to impact axial elongation through regulation of signaling pathways including canonical Wnt (*Wnt3a)* and RA (*Cyp26A1)* among others (Amin et al., 2016; Savory et al., 2009a; van Rooijen et al., 2012).

The further understanding of Cdx function in axial elongation requires a more comprehensive understanding of the landscape of the gene network within NMP cells. I therefore analyzed three

independent single cell RNA-sequencing (scRNAseq) datasets from three different mouse developmental stages published by other groups. This allowed examination of expression of *Cdx2* and putative direct targets *Sp8, Evx1, Zic3, Isl1* and *Nr2f1* in presumptive NMPs at the single cell level to gain insight into biological processes during axial elongation. Analysis of scRNA-sequence data from E7.0 mouse embryos, E8.0 caudal lateral epiblast cells or E9.5 posterior mouse embryos (Dias et al., 2020; Diaz-Cuadros et al., 2020; Scialdone et al., 2016) shows that *Sp8, Evx1, Zic3* and *Isl1* are co-expressed with *Cdx2* in cells that express the NMP markers *T/brachyury, Sox2,* and *Nkx1-2*, while *Nr2f1* is excluded from the same cells. The absence of *Isl1* in E9.5 scRNA-seq is probably due to the dataset having filtered genes for the PSM only. These observations lend further support of direct regulation of the expression of *Sp8, Evx1, Zic3, Isl1* and *Nr2f1* by Cdx2 at the single cell level of NMPs and their immediate descendants along E7.0, E8.0 and E9.5 pseudotime points.

Cdx regulates distinct processes through different targets for the emergence of post-occipital axial tissue

In vertebrates, the generation of PSM and coordinated body axis extension can be subdivided into three major phases that take place in the primitive streak and tail bud (Dubrulle and Pourquié, 2004). First, the balance between the differentiation and maintenance of NMPs in the progenitor state is necessary for the continuation of axis elongation at the proper rate. Second, the patterning and segmentation of cells migrating from the primitive streak or tail bud needs to occur properly. Finally, cells must be appropriately specified to the neural or paraxial mesoderm

lineage followed by the formation of somitic boundary. This series of events is regulated by an interactive cascade of transcription factors and signaling molecules, a number of which are downstream of Cdx2 (Savory et al., 2009a).

Cdx genes are required for the generation of the complete post-occipital morphogenesis. Previous research has well established that Cdx regulates the expression of *Hox* genes to influence vertebral AP patterning (Davidson et al., 2003; van Nes et al., 2006), but this cannot fully explain the phenotype of the Cdx mutants. Earlier work from our lab has shown the direct regulatory control by Cdx2 of *Wnt3a* and *Cyp26A1*, which are the important components of Wnt and RA signaling pathways, respectively, also influence axial elongation (Savory et al., 2009a), implicating Cdx function in the generation and maintenance of NMPs during mouse embryonic trunk and tail formation.

My present findings, together with prior work, suggests that Cdx contributes to the axial elongation via additional mechanisms, including regulation of the novel targets *Sp8, Evx1, Zic3, Isl1* and *Nr2f1* as described in this study. In addition to genes, considerable additional evidence supports a functional relationship between Cdx and these targets. For example, ablation of *Sp8* results in the loss of caudal somites, defects in neuropore closure and dysregulation of neural development due to a failure to maintain NMP cells (Bell et al., 2003; Kennedy et al., 2016); similar defects are observed in *Cdx* mutants (Amin et al., 2016; Bell et al., 2003; Kennedy et al., 2016; Savory et al., 2009a). *Sp8* is a mediator of FGF signaling, and *Sp8* exhibits a unique reciprocal induction loop with *Fgf8* believed to be required for the maintenance of NMP

progenitors (Sahara et al., 2007). Consistent with this, *Fgf8* expression is lost in Cdx mutants (Savory et al., 2009a), further supporting this relationship.

Evx1 is located 50kb 3' of the *Hoxa* cluster and is expressed in the posterior primitive streak from E6.5 (Bell et al., 2016). *Evx1* has been proposed to function in a regulatory network with *T/brachyury* to control AP cell fates and as a driver of axial elongation (Frith et al., 2018). Recent evidence also suggests that *Evx1* is an important downstream effector of the canonical Wnt signaling pathway and functions to maintain posterior identity by suppressing the expression of anterior genes (Bell et al., 2016).

Zic3 functions in the primitive streak to control the proliferation of paraxial mesoderm progenitor cells in the tail bud (Inoue et al., 2007). The loss of *Zic3* leads to defects in both gastrulation and neural tube closure, as well as a reduction in the total number of somites (Cast et al., 2012); again these defects recapitulate various aspects seen in Cdx mutants.

Isl1 is downstream of the canonical Wnt signaling pathway and induces hindlimbs and cloacal tissues. It also serves to transition NMPs into terminal differentiation pathways (Jurberg et al., 2013; Lin et al., 2007), and has been proposed to function in trunk-to-tail transition, which is also impacted in *Cdx*, *T/brachyury* and *Hox13* mutants (Amin et al., 2016; Sturgeon et al., 2011).

Finally, *Nr2f1* is expressed in more anterior tissues downstream of RA but it is excluded from the tail bud in the Cdx expression domain, and functions redundantly to control axis elongation

with *Nr2f2* (Béland and Lohnes, 2005; Vilhais-Neto et al., 2010). *Nr2f1* is also a key transcription factor for neural crest formation (Frith et al., 2018), another role of Cdx protein during embryogenesis (Sanchez-Ferras et al., 2016).

Cdx2 and chromosome configuration

Cells achieve differential gene expression by implementing gene regulatory networks which can be described as complex interplay between two components: the target gene and its regulators (Panigrahi and O'Malley, 2021). The regulators consist of both the *cis*-regulatory landscape, and *trans*-acting transcription factors as well as permissive or restrictive epigenetic components. As regards the latter, in addition to proximal and distal *cis*-regulatory elements (CRE), the 3-dimensional organization and spatial compartmentalization of loci plays essential roles in gene regulation (de Laat and Duboule, 2013). At the megabase scale, mammalian chromosomes are partitioned into active and inactive compartments, and these compartments can be subdivided into a set of self-interacting compartments called topologically associating domains (TADs) (Ulianov et al., 2016). The discovery of TADs was a very important concept towards obtaining a more comprehensive picture of the logic underlying gene regulation. For example, the HoxA and HoxD clusters were each found to lie at the junction of two TADs, which could possibly underlie the stepwise expression from 3' *Hox* genes to 5' neighbors during mouse limb outgrowth and axial patterning (Andrey et al., 2013; Darbellay and Duboule, 2016).

I noted that certain of the Cdx2 binding peaks identified in the ChIP-seq analysis are distal from the transcription start site (TSS) of the annotated target genes. For example, the Cdx2 binding peak for *Zic3* is ~120kb upstream of the TSS for this gene. However, it is possible that such Cdx2 binding sites and the relevant target gene(s) are located in the same TADs which infers that the Cdx proteins complexed to the distal CDRE interact more frequently with gene proximal elements to impact target gene expression. To address this issue, I analyzed Hi-C sequencing data for *Sp8, Evx1, Zic3, Isl1* and *Nr2f1* loci using published data from mouse embryonic stem cells (Dixon et al., 2012). The result from this analysis confirms that the Cdx2-bound regions and the corresponding Cdx target genes were located in the same TAD, including targets with distal CDREs, such as *Zic3*. I also found that all four murine *Hox* clusters are in TADs with strong Cdx2 binding peaks, which is consistent with the known regulation of *Hox* genes by Cdx members, and in agreement with previous findings (Neijts et al., 2016). It is possible that Cdx2 is able to interact with Condensin, CTCF and other TAD-related factors to form a sub-compartment on the chromatin, which facilitates transcriptional regulation of target genes factor activities more efficiently. While speculative, ChiP-seq data shows coincident genome occupation of Cdx, CTCF and Condensins at several target gene loci, consistent with their potential inteaction. Thus, these findings suggest a novel spatiotemporal control of Cdx2 as relates to on target gene regulation.

Future directions

My data has revealed key roles for Cdx in axial elongation and demonstrated that Cdx directly regulates the expression of 5 potential novel target genes, *Sp8, Evx1, Zic3, Isl1* and *Nr2f1*. These

results shed light on the key components of the complex genetic network behind axial elongation, yet there are still a number of questions that remain unanswered in regard to Cdx function in trunk and tail formation.

Following validation of putative target genes, it will be interesting to characterise their biological functions. This objective can be accomplished by complementation study using *Cdx1*$^{-/-}$*Cdx2*$^{+/-}$ heterozygous mice crossed with heterozygous mutants for each target gene to determine the phenotypes of mouse embryos or adult mice compared to the phenotypes of *Cdx1/2* double mutants. Such an approach could also be greatly aided by the gastruloids, a powerful tool that recapitulates the axial organization of post-implantation embryos and to study post-implantation embryonic development *in vitro.* Since this gastruloid field is still young and there exists some limitations of this model such as reproducibility (van den Brink and van Oudenaarden, 2021), it could still be used to obtain fundamental insights in the embryonic developmental processes such as axial elongation. For example, this *in vitro* culture allows us to observe the altered Cdx target gene expression patterns in WT or CRISPR edited Cdx null cells by fluorescent reporters in a high-throughput manner.

Secondly, there are several additional genes recovered in the RNA-seq analysis that are worth further investigation. For example, Sall4 is one of top genes in the RNA-seq list. Recent studies showed that Sall4 regulates the balance between maintenance and differentiation of NMPs into neurons and mesoderm lineage (Tahara et al., 2019). In *Sall4* conditional knockout embryos, NMPs lose bi-potency and favor the neural lineage, leading to a shorter posterior region of the mutants. Although *Sall4* may not be a direct Cdx target, it is possible that Cdx controls genes

that are implicated in axial elongation upstream of *Sall4*. Identification of such gene regulatory networks would expand on our understanding of the function of Cdx2 during development.

Another attractive direction would be the examination of the higher order chromosome architecture organization by Cdx2. Previous work in our lab suggests that Cdx2 regulates target genes through recruitment of Brg1-associated SWI-SNF chromatin remodeling activity (Nguyen et al., 2017b). Furthermore, Cdx2 has been suggested to act as a pioneer factor, which helps shape chromatin to become more accessible to regulation by secondary transcription factors (Kumar et al., 2019). These findings indicate that Cdx2 is able to regulate target genes via spatiotemporal manner.

While I was using Hi-C seq to show that the Cdx2 binding sites of *Zic3* were located in the same topological associating domain with *Zic3* – and therefore could still interact frequently within the TAD - I found in some cases Cdx2 occupancy co-localized with peaks of Condensin II and CCCTC-binding factor (CTCF). Condensin II complexes regulate large-scale interactions and spatial partitioning of chromosomes during interphase (Rosin et al., 2018), and CTCF functions as a transcriptional regulator that recruits other transcription factors while bound to chromatin domain boundaries. Notably, this colocalization also happens in the human intestine, another Cdx-dependent tissue. These findings suggest another whole mechanism for Cdx2 function in regulating the expression of targets. In this case, we could use immunoprecipitation and mass spectroscopy to determine if CTCF or Condensin II will co-immunoprecipitate with Cdx from mouse tail buds.

Summary

This work revealed key roles for Cdx in axial elongation through the direct regulation of 5 novel target genes, *Sp8, Evx1, Zic3, Isl1* and *Nr2f1*. We have generated a conditional *Cdx1/2* double mutant mouse model in combination with a transgenic mouse line *Pax2*-GFP to circumvent Cdx function overlap and peri-implantation lethality of Cdx2 null mutants in the mouse and enrich for NMP cells for analysis. We showed that Cdx was able to induce the expression of *Sp8, Evx1, Zic3, Isl1* and repress expression of *Nf2f1* via binding to conserved elements within the proximal promoter region or distal enhancer elements which harbour at least one CDRE. These results further integrate new roles for Cdx in a complex genetic network for the maintenance of axial progenitors and their subsequent differentiation during axial extension.

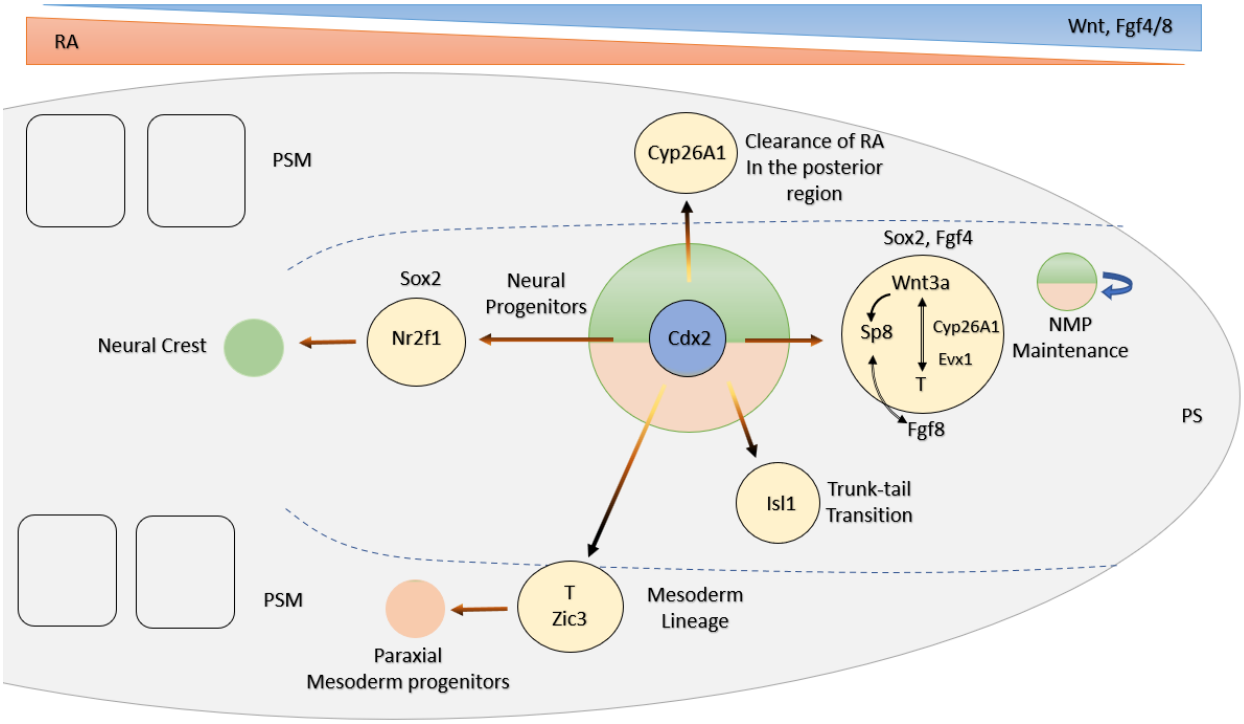


Figure 14. *Cdx* as a central regulator of the axial elongation process. Schematic of Cdx-dependent gene regulation in NMPs and in axial elongation.

REFERENCES

Abu-Abed, S., Dollé, P., Metzger, D., Beckett, B., Chambon, P., Petkovich, M., 2001. The retinoic acid-metabolizing enzyme, CYP26A1, is essential for normal hindbrain patterning, vertebral identity, and development of posterior structures. Genes Dev 15, 226-240.

Acloque, H., Adams, M.S., Fishwick, K., Bronner-Fraser, M., Nieto, M.A., 2009. Epithelial-mesenchymal transitions: the importance of changing cell state in development and disease. J Clin Invest 119, 1438-1449.

Amin, S., Neijts, R., Simmini, S., van Rooijen, C., Tan, S.C., Kester, L., van Oudenaarden, A., Creyghton, M.P., Deschamps, J., 2016. Cdx and T Brachyury Co-activate Growth Signaling in the Embryonic Axial Progenitor Niche. Cell Reports 17, 3165-3177.

Andrey, G., Montavon, T., Mascrez, B., Gonzalez, F., Noordermeer, D., Leleu, M., Trono, D., Spitz, F., Duboule, D., 2013. A switch between topological domains underlies HoxD genes collinearity in mouse limbs. Science (New York, N.Y.) 340, 1234167.

Arnold, S.J., Robertson, E.J., 2009. Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. Nature Reviews Molecular Cell Biology 10, 91-103.

Aulehla, A., Wehrle, C., Brand-Saberi, B., Kemler, R., Gossler, A., Kanzler, B., Herrmann, B.G., 2003. Wnt3a Plays a Major Role in the Segmentation Clock Controlling Somitogenesis. Developmental Cell 4, 395-406.

Beck, F., 2002. Homeobox genes in gut development. Gut 51, 450-454.

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. Developmental dynamics : an official publication of the American Association of Anatomists 204, 219-227.

Beddington, R.S., Robertson, E.J., 1999. Axis development and early asymmetry in mammals. Cell 96, 195-209.

Béland, M., Lohnes, D., 2005. Chicken Ovalbumin Upstream Promoter-Transcription Factor Members Repress Retinoic Acid-induced Cdx1 Expression. Journal of Biological Chemistry 280, 13858-13862.

Bell, C.C., Amaral, P.P., Kalsbeek, A., Magor, G.W., Gillinder, K.R., Tangermann, P., di Lisio, L., Cheetham, S.W., Gruhl, F., Frith, J., Tallack, M.R., Ru, K.-L., Crawford, J., Mattick, J.S., Dinger, M.E., Perkins, A.C., 2016. The Evx1/Evx1as gene locus regulates anterior-posterior patterning during gastrulation. Scientific Reports 6, 26657.

Bell, S.M., Schreiner, C.M., Waclaw, R.R., Campbell, K., Potter, S.S., Scott, W.J., 2003. Sp8 is crucial for limb outgrowth and neuropore closure. 100, 12195-12200.

Böhmer, C., Werneburg, I., 2017. Deep time perspective on turtle neck evolution: chasing the Hox code by vertebral morphology. Sci Rep 7, 8939.
Boulet, A.M., Capecchi, M.R., 2012. Signaling by FGF4 and FGF8 is required for axial elongation of the mouse embryo. Developmental biology 371, 235-245.
Brooke-Bisschop, T., Savory, J.G.A., Foley, T., Ringuette, R., Lohnes, D., 2017. Essential roles for Cdx in murine primitive hematopoiesis. Developmental Biology 422, 115-124.
Cambray, N., Wilson, V., 2002. Axial progenitors with extensive potency are localised to the mouse chordoneural hinge. 129, 4855-4866.
Cambray, N., Wilson, V., 2007. Two distinct sources for a population of maturing axial progenitors. Development 134, 2829-2840.
Cast, A.E., Gao, C., Amack, J.D., Ware, S.M., 2012. An essential and highly conserved role for Zic3 in left-right patterning, gastrulation and convergent extension morphogenesis. Developmental biology 364, 22-31.
Charité, J., de Graaff, W., Consten, D., Reijnen, M.J., Korving, J., Deschamps, J., 1998. Transducing positional information to the Hox genes: critical interaction of cdx gene products with position-sensitive regulatory elements. Development 125, 4349-4358.
Chase, A., Reiter, A., Burci, L., Cazzaniga, G., Biondi, A., Pickard, J., Roberts, I.A., Goldman, J.M., Cross, N.C., 1999. Fusion of ETV6 to the caudal-related homeobox gene CDX2 in acute myeloid leukemia with the t(12;13)(p13;q12). Blood 93, 1025-1031.
Chawengsaksophak, K., de Graaff, W., Rossant, J., Deschamps, J., Beck, F., 2004. Cdx2 is essential for axial elongation in mouse development. 101, 7641-7645.
Chawengsaksophak, K., James, R., Hammond, V.E., Köntgen, F., Beck, F., 1997. Homeosis and intestinal tumours in Cdx2 mutant mice. Nature 386, 84-87.
Ciruna, B., Rossant, J., 2001. FGF Signaling Regulates Mesoderm Cell Fate Specification and Morphogenetic Movement at the Primitive Streak. Developmental Cell 1, 37-49.
Clevers, H., Nusse, R., 2012. Wnt/β-Catenin Signaling and Disease. Cell 149, 1192-1205.
Croce, J.C., McClay, D.R., 2008. Evolution of the Wnt pathways. Methods Mol Biol 469, 3-18.
Crossley, P.H., Martin, G.R., 1995. The mouse Fgf8 gene encodes a family of polypeptides and is expressed in regions that direct outgrowth and patterning in the developing embryo. Development 121, 439-451.
Darbellay, F., Duboule, D., 2016. Topological Domains, Metagenes, and the Emergence of Pleiotropic Regulations at Hox Loci. Current topics in developmental biology 116, 299-314.

Davidson, A.J., Ernst, P., Wang, Y., Dekens, M.P.S., Kingsley, P.D., Palis, J., Korsmeyer, S.J., Daley, G.Q., Zon, L.I., 2003. cdx4 mutants fail to specify blood progenitors and can be rescued by multiple hox genes. Nature 425, 300-306.
Davidson, A.J., Zon, L.I., 2006. The caudal-related homeobox genes cdx1a and cdx4 act redundantly to regulate hox gene expression and the formation of putative hematopoietic stem cells during zebrafish embryogenesis. Dev Biol 292, 506-518.
de Laat, W., Duboule, D., 2013. Topology of mammalian developmental enhancers and their regulatory landscapes. Nature 502, 499-506.
Dearolf, C.R., Topol, J., Parker, C.S., 1989. The caudal gene product is a direct activator of fushi tarazu transcription during Drosophila embryogenesis. Nature 341, 340-343.
del Corral, R.D., Olivera-Martinez, I., Goriely, A., Gale, E., Maden, M., Storey, K., 2003. Opposing FGF and Retinoid Pathways Control Ventral Neural Pattern, Neuronal Differentiation, and Segmentation during Body Axis Extension. Neuron 40, 65-79.
Dias, A., Lozovska, A., Wymeersch, F.J., Nóvoa, A., Binagui-Casas, A., Sobral, D., Martins, G.G., Wilson, V., Mallo, M., 2020. A Tgfbr1/Snai1-dependent developmental module at the core of vertebrate axial elongation. eLife 9, e56615.
Diaz-Cuadros, M., Wagner, D.E., Budjan, C., Hubaud, A., Tarazona, O.A., Donelly, S., Michaut, A., Al Tanoury, Z., Yoshioka-Kobayashi, K., Niino, Y., Kageyama, R., Miyawaki, A., Touboul, J., Pourquié, O., 2020. In vitro characterization of the human segmentation clock. Nature 580, 113-118.
Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W., Diao, Y., Liang, J., Zhao, H., Lobanenkov, V.V., Ecker, J.R., Thomson, J.A., Ren, B., 2015. Chromatin architecture reorganization during stem cell differentiation. Nature 518, 331-336.
Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., Ren, B., 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485, 376-380.
Dorey, K., Amaya, E., 2010. FGF signalling: diverse roles during early vertebrate embryogenesis. Development (Cambridge, England) 137, 3731-3742.
Duboule, D., 1998. Vertebrate hox gene regulation: clustering and/or colinearity? Current opinion in genetics & development 8, 514-518.
Duboule, D., Dollé, P., 1989. The structural and functional organization of the murine HOX gene family resembles that of Drosophila homeotic genes. The EMBO journal 8, 1497-1505.
Dubrulle, J., Pourquié, O., 2004. Coupling segmentation to axis formation. Development 131, 5783-5793.
Duester, G., 2008. Retinoic Acid Synthesis and Signaling during Early Organogenesis. Cell 134, 921-931.

Dunty, W.C., Jr, Biris, K.K., Chalamalasetty, R.B., Taketo, M.M., Lewandoski, M., Yamaguchi, T.P., 2008. Wnt3a/β-catenin signaling controls posterior body development by coordinating mesoderm formation and segmentation. Development 135, 85-94.
Eden, E., Lipson, D., Yogev, S., Yakhini, Z., 2007. Discovering motifs in ranked lists of DNA sequences. PLoS computational biology 3, e39.
Eden, E., Navon, R., Steinfeld, I., Lipson, D., Yakhini, Z., 2009. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC bioinformatics 10, 48.
Evans, A.L., Faial, T., Gilchrist, M.J., Down, T., Vallier, L., Pedersen, R.A., Wardle, F.C., Smith, J.C., 2012. Genomic Targets of Brachyury (T) in Differentiating Mouse Embryonic Stem Cells. PLOS ONE 7, e33346.
Feldman, B., Poueymirou, W., Papaioannou, V.E., DeChiara, T.M., Goldfarb, M., 1995. Requirement of FGF-4 for postimplantation mouse development. Science (New York, N.Y.) 267, 246-249.
Fleming, T.P., 1987. A quantitative analysis of cell allocation to trophectoderm and inner cell mass in the mouse blastocyst. Developmental Biology 119, 520-531.
Foley, T.E., Hess, B., Savory, J.G.A., Ringuette, R., Lohnes, D., 2019. Role of Cdx factors in early mesodermal fate decisions. Development 146, dev170498.
Frith, T.J.R., Granata, I., Wind, M., Stout, E., Thompson, O., Neumann, K., Stavish, D., Heath, P.R., Ortmann, D., Hackland, J.O.S., Anastassiadis, K., Gouti, M., Briscoe, J., Wilson, V., Johnson, S.L., Placzek, M., Guarracino, M.R., Andrews, P.W., Tsakiridis, A., 2018. Human axial progenitors generate trunk neural crest cells in vitro. eLife 7, e35786.
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.
García-García, M.J., Shibata, M., Anderson, K.V., 2008. Chato, a KRAB zinc-finger protein, regulates convergent extension in the mouse embryo. 135, 3053-3062.
Gardner, R.L., 1978. The relationship between cell lineage and differentiation in the early mouse embryo. Results and problems in cell differentiation 9, 205-241.
Gentsch, G.E., Owens, N.D., Martin, S.R., Piccinelli, P., Faial, T., Trotter, M.W., Gilchrist, M.J., Smith, J.C., 2013. In vivo T-box transcription factor profiling reveals joint regulation of embryonic neuromesodermal bipotency. Cell Rep 4, 1185-1196.
Ghyselinck, N.B., Duester, G., 2019. Retinoic acid signaling pathways. Development 146.
Gordon, M.D., Nusse, R., 2006. Wnt signaling: multiple pathways, multiple receptors, and multiple transcription factors. The Journal of biological chemistry 281, 22429-22433.
Grainger, S., Lam, J., Savory, J.G., Mears, A.J., Rijli, F.M., Lohnes, D., 2012. Cdx regulates Dll1 in multiple lineages. Dev Biol 361, 1-11.

Greco, T.L., Takada, S., Newhouse, M.M., McMahon, J.A., McMahon, A.P., Camper, S.A., 1996. Analysis of the vestigial tail mutation demonstrates that Wnt-3a gene dosage regulates mouse axial development. 10, 313-324.
Guazzi, S., Lonigro, R., Pintonello, L., Boncinelli, E., Di Lauro, R., Mavilio, F., 1994. The thyroid transcription factor-1 gene is a candidate target for regulation by Hox proteins. The EMBO journal 13, 3339-3347.
Henrique, D., Abranches, E., Verrier, L., Storey, K.G., 2015. Neuromesodermal progenitors and the making of the spinal cord. Development 142, 2864-2875.
Herrmann, B.G., Kispert, A., 1994. The T genes in embryogenesis. Trends in Genetics 10, 280-286.
Herrmann, B.G., Labeit, S., Poustka, A., King, T.R., Lehrach, H., 1990. Cloning of the T gene required in mesoderm formation in the mouse. Nature 343, 617-622.
Hikasa, H., Sokol, S.Y., 2013. Wnt signaling in vertebrate axis specification. Cold Spring Harb Perspect Biol 5, a007955-a007955.
Houle, M., Allan, D., Lohnes, D., 2003. Cdx homeodomain proteins in vertebral patterning, Advances in Developmental Biology and Biochemistry. Elsevier, pp. 69-105.
Iimura, T., Pourquié, O., 2007. Hox genes in time and space during vertebrate body formation. Development, growth & differentiation 49, 265-275.
Inman, K.E., Downs, K.M., 2006. Localization of Brachyury (T) in embryonic and extraembryonic tissues during mouse gastrulation. Gene expression patterns : GEP 6, 783-793.
Inoue, T., Ota, M., Mikoshiba, K., Aruga, J., 2007. Zic2 and Zic3 synergistically control neurulation and segmentation of paraxial mesoderm in mouse embryo. Developmental Biology 306, 669-684.
Iwasaki, O., Tanizawa, H., Kim, K.-D., Kossenkov, A., Nacarelli, T., Tashiro, S., Majumdar, S., Showe, L.C., Zhang, R., Noma, K.-i., 2019. Involvement of condensin in cellular senescence through gene regulation and compartmental reorganization. Nature Communications 10, 5688.
Janesick, A., Nguyen, T.T., Aisaki, K., Igarashi, K., Kitajima, S., Chandraratna, R.A., Kanno, J., Blumberg, B., 2014. Active repression by RARγ signaling is required for vertebrate axial elongation. Development 141, 2260-2270.
Jurberg, Arnon D., Aires, R., Varela-Lasheras, I., Nóvoa, A., Mallo, M., 2013. Switching Axial Progenitors from Producing Trunk to Tail Tissues in Vertebrate Embryos. Developmental Cell 25, 451-462.
Kennedy, M.W., Chalamalasetty, R.B., Thomas, S., Garriock, R.J., Jailwala, P., Yamaguchi, T.P., 2016. Sp5 and Sp8 recruit β-catenin and Tcf1-Lef1 to select enhancers to activate Wnt target gene transcription. 113, 3545-3550.

Kessel, M., Gruss, P., 1991. Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. Cell 67, 89-104.
Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., Salzberg, S.L., 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biology 14, R36.
Kohn, A.D., Moon, R.T., 2005. Wnt and calcium signaling: beta-catenin-independent pathways. Cell calcium 38, 439-446.
Komiya, Y., Habas, R., 2008. Wnt signal transduction pathways. Organogenesis 4, 68-75.
Korswagen, H.C., 2002. Canonical and non-canonical Wnt signaling pathways in Caenorhabditis elegans: variations on a common signaling theme. BioEssays : news and reviews in molecular, cellular and developmental biology 24, 801-810.
Krumlauf, R., 1994. Hox genes in vertebrate development. Cell 78, 191-201.
Kumar, N., Tsai, Y.-H., Chen, L., Zhou, A., Banerjee, K.K., Saxena, M., Huang, S., Toke, N.H., Xing, J., Shivdasani, R.A., Spence, J.R., Verzi, M.P., 2019. The lineage-specific transcription factor CDX2 navigates dynamic chromatin to control distinct stages of intestine development. Development 146.
Levine, M., Harding, K., Wedeen, C., Doyle, H., Hoey, T., Radomska, H., 1985. Expression of the homeo box gene family in Drosophila. Cold Spring Harbor symposia on quantitative biology 50, 209-222.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., Subgroup, G.P.D.P., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078-2079.
Lin, L., Cui, L., Zhou, W., Dufort, D., Zhang, X., Cai, C.-L., Bu, L., Yang, L., Martin, J., Kemler, R., Rosenfeld, M.G., Chen, J., Evans, S.M., 2007. β-Catenin directly regulates Islet1 expression in cardiovascular progenitors and is required for multiple aspects of cardiogenesis. 104, 9313-9318.
Liu, P., Wakamiya, M., Shea, M.J., Albrecht, U., Behringer, R.R., Bradley, A., 1999. Requirement for Wnt3 in vertebrate axis formation. Nat Genet 22, 361-365.
Lohnes, D., 2003. The Cdx1 homeodomain protein: an integrator of posterior signaling in the mouse. 25, 971-980.
Macdonald, P.M., Struhl, G., 1986. A molecular gradient in early Drosophila embryos and its role in specifying the body pattern. Nature 324, 537-545.
Martin, B.L., Kimelman, D., 2010. Brachyury establishes the embryonic mesodermal progenitor niche. 24, 2778-2783.
Martin, B.L., Kimelman, D., 2012. Canonical Wnt signaling dynamically controls multiple stem cell fate decisions during vertebrate body formation. Developmental cell 22, 223-232.

Mathis, L., Nicolas, J.F., 2000. Different clonal dispersion in the rostral and caudal mouse central nervous system. 127, 1277-1290.
Mazzoni, E.O., Mahony, S., Peljto, M., Patel, T., Thornton, S.R., McCuine, S., Reeder, C., Boyer, L.A., Young, R.A., Gifford, D.K., Wichterle, H., 2013. Saltatory remodeling of Hox chromatin in response to rostrocaudal patterning signals. Nature Neuroscience 16, 1191-1198.
Meyer, B.I., Gruss, P., 1993. Mouse Cdx-1 expression during gastrulation. Development 117, 191-203.
Morkel, M., Huelsken, J., Wakamiya, M., Ding, J., van de Wetering, M., Clevers, H., Taketo, M.M., Behringer, R.R., Shen, M.M., Birchmeier, W., 2003. β-Catenin regulates Cripto- and Wnt3-dependent gene expression programs in mouse axis and mesoderm formation. Development 130, 6283-6294.
Nagy, A., Gertsenstein, M., Vintersten, K., Behringer, R., 2003. Manipulating the Mouse Embryo: A Laboratory Manual. Cold Spring Harbor Laboratory Press.
Naiche, L.A., Holder, N., Lewandoski, M., 2011. FGF4 and FGF8 comprise the wavefront activity that controls somitogenesis. Proc Natl Acad Sci U S A 108, 4018-4023.
Neijts, R., Amin, S., van Rooijen, C., Tan, S., Creyghton, M.P., de Laat, W., Deschamps, J., 2016. Polarized regulatory landscape and Wnt responsiveness underlie Hox activation in embryos. Genes Dev 30, 1937-1942.
Nguyen, T.T., Savory, J.G.A., Brooke-Bisschop, T., Ringuette, R., Foley, T., Hess, B.L., Mulatz, K.J., Trinkle-Mulcahy, L., Lohnes, D., 2017a. Cdx2 Regulates Gene Expression through Recruitment of Brg1-associated Switch-Sucrose Non-fermentable (SWI-SNF) Chromatin Remodeling Activity. The Journal of biological chemistry 292, 3389-3399.
Nguyen, T.T., Savory, J.G.A., Brooke-Bisschop, T., Ringuette, R., Foley, T., Hess, B.L., Mulatz, K.J., Trinkle-Mulcahy, L., Lohnes, D., 2017b. Cdx2 Regulates Gene Expression through Recruitment of Brg1-associated Switch-Sucrose Non-fermentable (SWI-SNF) Chromatin Remodeling Activity. The Journal of biological chemistry 292, 3389-3399.
Niederreither, K., Subbarayan, V., Dollé, P., Chambon, P., 1999. Embryonic retinoic acid synthesis is essential for early mouse post-implantation development. Nature Genetics 21, 444-448.
Nishisho, I., Nakamura, Y., Miyoshi, Y., Miki, Y., Ando, H., Horii, A., Koyama, K., Utsunomiya, J., Baba, S., Hedge, P., 1991. Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. Science (New York, N.Y.) 253, 665-669.
Niswander, L., Martin, G.R., 1992. Fgf-4 expression during gastrulation, myogenesis, limb and tooth development in the mouse. Development 114, 755-768.
Nowotschin, S., Ferrer-Vaquer, A., Concepcion, D., Papaioannou, V.E., Hadjantonakis, A.-K., 2012. Interaction of Wnt3a, Msgn1 and Tbx6 in neural versus paraxial mesoderm lineage commitment and paraxial mesoderm differentiation in the mouse embryo. Developmental Biology 367, 1-14.

Olivera-Martinez, I., Harada, H., Halley, P.A., Storey, K.G., 2012. Loss of FGF-dependent mesoderm identity and rise of endogenous retinoid signalling determine cessation of body axis elongation. PLoS Biol 10, e1001415-e1001415.
Ornitz, D.M., 2000. FGFs, heparan sulfate and FGFRs: complex interactions essential for development. BioEssays : news and reviews in molecular, cellular and developmental biology 22, 108-112.
Panigrahi, A., O'Malley, B.W., 2021. Mechanisms of enhancer action: the known and the unknown. Genome Biology 22, 108.
Partanen, J., Schwartz, L., Rossant, J., 1998. Opposite phenotypes of hypomorphic and Y766 phosphorylation site mutations reveal a function for Fgfr1 in anteroposterior patterning of mouse embryos. 12, 2332-2344.
Pineault, N., Helgason, C.D., Lawrence, H.J., Humphries, R.K., 2002. Differential expression of Hox, Meis1, and Pbx1 genes in primitive cells throughout murine hematopoietic ontogeny. Experimental Hematology 30, 49-57.
Prajapati, R.S., Mitter, R., Vezzaro, A., Ish-Horowicz, D., 2020. Greb1 is required for axial elongation and segmentation in vertebrate embryos. Biology Open 9.
Qiu, Y., Pereira, F.A., DeMayo, F.J., Lydon, J.P., Tsai, S.Y., Tsai, M.-J., 1997. Null mutation of mCOUP-TFI results in defects in morphogenesis of the glossopharyngeal ganglion, axonal projection, and arborization. 11, 1925-1937.
Quinlan, A.R., 2014. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. 47, 11.12.11-11.12.34.
Rappolee, D.A., Basilico, C., Patel, Y., Werb, Z., 1994. Expression and function of FGF-4 in peri-implantation development in mouse embryos. Development 120, 2259-2269.
Rhinn, M., Dollé, P., 2012. Retinoic acid signalling during development. 139, 843-858.
Rivera-Pérez, J.A., Magnuson, T., 2005. Primitive streak formation in mice is preceded by localized activation of Brachyury and Wnt3. Dev Biol 288, 363-371.
Rodrigo Albors, A., Halley, P.A., Storey, K.G., 2018. Lineage tracing of axial progenitors using Nkx1-2CreERT2 mice defines their trunk and tail contributions. Development 145.
Rosin, L.F., Nguyen, S.C., Joyce, E.F., 2018. Condensin II drives large-scale folding and spatial partitioning of interphase chromosomes in Drosophila nuclei. PLOS Genetics 14, e1007393.
Sahara, S., Kawakami, Y., Izpisua Belmonte, J.C., O'Leary, D.D.M., 2007. Sp8 exhibits reciprocal induction with Fgf8 but has an opposing effect on anterior-posterior cortical area patterning. Neural Development 2, 10.

Sakai, Y., Meno, C., Fujii, H., Nishino, J., Shiratori, H., Saijoh, Y., Rossant, J., Hamada, H., 2001. The retinoic acid-inactivating enzyme CYP26 is essential for establishing an uneven distribution of retinoic acid along the anterio-posterior axis within the mouse embryo. Genes Dev 15, 213-225.
Sanchez-Ferras, O., Bernas, G., Farnos, O., Touré, A.M., Souchkova, O., Pilon, N., 2016. A direct role for murine Cdx proteins in the trunk neural crest gene regulatory network. Development 143, 1363-1374.
Savory, J.G., Mansfield, M., St Louis, C., Lohnes, D., 2011a. Cdx4 is a Cdx2 target gene. Mech Dev 128, 41-48.
Savory, J.G.A., Bouchard, N., Pierre, V., Rijli, F.M., De Repentigny, Y., Kothary, R., Lohnes, D., 2009a. Cdx2 regulation of posterior development through non-Hox targets. 136, 4099-4110.
Savory, J.G.A., Mansfield, M., Rijli, F.M., Lohnes, D., 2011b. Cdx mediates neural tube closure through transcriptional regulation of the planar cell polarity gene. Development 138, 1361.
Savory, J.G.A., Pilon, N., Grainger, S., Sylvestre, J.-R., Béland, M., Houle, M., Oh, K., Lohnes, D., 2009b. Cdx1 and Cdx2 are functionally equivalent in vertebral patterning. Developmental Biology 330, 114-122.
Schefe, J.H., Lehmann, K.E., Buschmann, I.R., Unger, T., Funke-Kaiser, H., 2006. Quantitative real-time RT-PCR data analysis: current concepts and the novel "gene expression's CT difference" formula. Journal of molecular medicine (Berlin, Germany) 84, 901-910.
Scialdone, A., Tanaka, Y., Jawaid, W., Moignard, V., Wilson, N.K., Macaulay, I.C., Marioni, J.C., Göttgens, B., 2016. Resolving early mesoderm diversification through single-cell expression profiling. Nature 535, 289-293.
Sharma, R., Shafer, M.E.R., Bareke, E., Tremblay, M., Majewski, J., Bouchard, M., 2017. Bmp signaling maintains a mesoderm progenitor cell state in the mouse tailbud. Development 144, 2982.
Simons, M., Mlodzik, M., 2008. Planar cell polarity signaling: from fly development to human disease. Annual review of genetics 42, 517-540.
Sirbu, I.O., Duester, G., 2006. Retinoic-acid signalling in node ectoderm and posterior neural plate directs left-right patterning of somitic mesoderm. Nature cell biology 8, 271-277.
Stott, D., Kispert, A., Herrmann, B.G., 1993. Rescue of the tail defect of Brachyury mice. Genes Dev 7, 197-203.
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.
Sturgeon, K., Kaneko, T., Biemann, M., Gauthier, A., Chawengsaksophak, K., Cordes, S.P., 2011. Cdx1 refines positional identity of the vertebrate hindbrain by directly repressing Mafb expression. Development 138, 65-74.

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.
Sun, X., Meyers, E.N., Lewandoski, M., Martin, G.R., 1999. Targeted disruption of Fgf8 causes failure of cell migration in the gastrulating mouse embryo. Genes Dev 13, 1834-1846.
Tahara, N., Kawakami, H., Chen, K.Q., Anderson, A., Yamashita Peterson, M., Gong, W., Shah, P., Hayashi, S., Nishinakamura, R., Nakagawa, Y., Garry, D.J., Kawakami, Y., 2019. Sall4 regulates neuromesodermal progenitors and their descendants during body elongation in mouse embryos. Development 146.
Takada, S., Stark, K.L., Shea, M.J., Vassileva, G., McMahon, J.A., McMahon, A.P., 1994. Wnt-3a regulates somite and tailbud formation in the mouse embryo. Genes Dev 8, 174-189.
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.L., Behringer, R.R., 1997. Mouse gastrulation: the formation of a mammalian body plan. Mechanisms of Development 68, 3-25.
Taylor, J.K., Levy, T., Suh, E.R., Traber, P.G., 1997. Activation of enhancer elements by the homeobox gene Cdx2 is cell line specific. Nucleic Acids Res 25, 2293-2300.
Teven, C.M., Farina, E.M., Rivas, J., Reid, R.R., 2014. Fibroblast growth factor (FGF) signaling in development and skeletal diseases. Genes Dis 1, 199-213.
Tortelote, G.G., Hernández-Hernández, J.M., Quaresma, A.J., Nickerson, J.A., Imbalzano, A.N., Rivera-Pérez, J.A., 2013. Wnt3 function in the epiblast is required for the maintenance but not the initiation of gastrulation in mice. Dev Biol 374, 164-173.
Tsakiridis, A., Huang, Y., Blin, G., Skylaki, S., Wymeersch, F., Osorno, R., Economou, C., Karagianni, E., Zhao, S., Lowell, S., Wilson, V., 2014. Distinct Wnt-driven primitive streak-like populations reflect in vivo lineage precursors. Development 141, 1209-1221.
Tzouanacou, E., Wegener, A., Wymeersch, F.J., Wilson, V., Nicolas, J.F., 2009. Redefining the progression of lineage segregations during mammalian embryogenesis by clonal analysis. Dev Cell 17, 365-376.
Ulianov, S.V., Khrameeva, E.E., Gavrilov, A.A., Flyamer, I.M., Kos, P., Mikhaleva, E.A., Penin, A.A., Logacheva, M.D., Imakaev, M.V., Chertovich, A., Gelfand, M.S., Shevelyov, Y.Y., Razin, S.V., 2016. Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains. Genome Res 26, 70-84.
van Amerongen, R., Nusse, R., 2009. Towards an integrated view of Wnt signaling in development. 136, 3205-3214.

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. 129, 2181-2193.
van den Brink, S.C., van Oudenaarden, A., 2021. 3D gastruloids: a novel frontier in stem cell-based in vitro modeling of mammalian gastrulation. Trends in Cell Biology 31, 747-759.
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 Rooijen, C., Simmini, S., Bialecka, M., Neijts, R., van de Ven, C., Beck, F., Deschamps, J., 2012. Evolutionarily conserved requirement of Cdx for post-occipital tissue emergence. Development 139, 2576-2583.
Vermot, J., Gallego Llamas, J., Fraulob, V., Niederreither, K., Chambon, P., Dollé, P., 2005. Retinoic acid controls the bilateral symmetry of somite formation in the mouse embryo. Science (New York, N.Y.) 308, 563-566.
Vietri Rudan, M., Barrington, C., Henderson, S., Ernst, C., Odom, D.T., Tanay, A., Hadjur, S., 2015. Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. Cell reports 10, 1297-1309.
Vilhais-Neto, G.C., Maruhashi, M., Smith, K.T., Vasseur-Cognet, M., Peterson, A.S., Workman, J.L., Pourquié, O., 2010. Rere controls retinoic acid signalling and somite bilateral symmetry. Nature 463, 953-957.
Wahl, M.B., Deng, C., Lewandoski, M., Pourquié, O., 2007. FGF signaling acts upstream of the NOTCH and WNT signaling pathways to control segmentation clock oscillations in mouse somitogenesis. Development 134, 4033.
Wallingford, J.B., 2006. Planar cell polarity, ciliogenesis and neural tube defects. Human Molecular Genetics 15, R227-R234.
Wallingford, J.B., Habas, R., 2005. The developmental biology of Dishevelled: an enigmatic protein governing cell fate and cell polarity. 132, 4421-4436.
Wang, Y., Yabuuchi, A., McKinney-Freeman, S., Ducharme, D.M.K., Ray, M.K., Chawengsaksophak, K., Archer, T.K., Daley, G.Q., 2008. Cdx Gene Deficiency Compromises Embryonic Hematopoiesis in the Mouse. Proc Natl Acad Sci U S A 105, 7756-7761.
Wellik, D.M., 2007. Hox patterning of the vertebrate axial skeleton. Developmental dynamics : an official publication of the American Association of Anatomists 236, 2454-2463.
Wharton, K.A., 2003. Runnin' with the Dvl: Proteins That Associate with Dsh/Dvl and Their Significance to Wnt Signal Transduction. Developmental Biology 253, 1-17.

Wolpert, L., 2019. Principles of development.
Wymeersch, F.J., Huang, Y., Blin, G., Cambray, N., Wilkie, R., Wong, F.C.K., Wilson, V., 2016. Position-dependent plasticity of distinct progenitor types in the primitive streak. eLife 5, e10042.
Xu, X., Li, C., Takahashi, K., Slavkin, H.C., Shum, L., Deng, C.-X., 1999. Murine Fibroblast Growth Factor Receptor 1α Isoforms Mediate Node Regression and Are Essential for Posterior Mesoderm Development. Developmental Biology 208, 293-306.
Yamaguchi, T.P., Takada, S., Yoshikawa, Y., Wu, N., McMahon, A.P., 1999. T (Brachyury) is a direct target of Wnt3a during paraxial mesoderm specification. Genes Dev 13, 3185-3190.
Yoshikawa, Y., Fujimori, T., McMahon, A.P., Takada, S., 1997. Evidence that absence of Wnt-3a signaling promotes neuralization instead of paraxial mesoderm development in the mouse. Dev Biol 183, 234-242.
Young, T., Rowland, J.E., van de Ven, C., Bialecka, M., Novoa, A., Carapuco, M., van Nes, J., de Graaff, W., Duluc, I., Freund, J.-N., Beck, F., Mallo, M., Deschamps, J., 2009. Cdx and Hox Genes Differentially Regulate Posterior Axial Growth in Mammalian Embryos. Developmental Cell 17, 516-526.
Zhang, X., Ibrahimi, O.A., Olsen, S.K., Umemori, H., Mohammadi, M., Ornitz, D.M., 2006. Receptor Specificity of the Fibroblast Growth Factor Family: The complete mammalian FGF family. Journal of Biological Chemistry 281, 15694-15700.