

**Mouse Hox and Cdx genes:
Early regulation and involvement in
embryonic axial extension**

Roel Neijts

Utrecht, 2017

The work described in this thesis was performed at the Hubrecht Institute for Developmental Biology and Stem Cell Research (the Royal Netherlands Academy of Arts and Sciences, KNAW) within the framework of the research school Cancer Stem cells & Developmental biology (CS&D), which is part of the Utrecht Graduate School of Life Sciences (Utrecht University).

Cover: 'From embryo to 4C-seq peak', by Roel Neijts and Carina van Rooijen.

Layout: Roel Neijts

Printing: Gildeprint, Enschede

ISBN: 978-94-6233-5097

Copyright © 2017 by Roel Neijts. All rights reserved. No part of this book may be reproduced, stored in a retrieval system or transmitted in any form or by any means, without prior permission of the author.

Mouse Hox and Cdx genes:

Early regulation and involvement in embryonic axial extension

Hox- en Cdx-genen in muizen:
hun vroege regulatie en hun invloed op embryonale lichaamsas-generatie
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector
magnificus, prof.dr. G.J. van der Zwaan, ingevolge het besluit van het college van promoties
in het openbaar te verdedigen op dinsdag 24 januari 2017 des middags te 12.45 uur

door

Roel Neijts

geboren op 14 oktober 1985 te Bakel en Milheeze

Promoter: prof. dr. H.C. Clevers
Co-promoter: dr. J.T. Deschamps

The research described in this thesis was supported by a grant from the Dutch government to the Netherlands Institute for Regenerative Medicine (NIRM, grant FES0908).

Table of contents

Summary of the thesis		6
Chapter 1	<i>Introduction part A</i>	8
	Region-specific regulation of posterior axial elongation during vertebrate embryogenesis	
	<i>Introduction part B</i>	30
	At the start of colinear Hox gene expression: <i>cis</i> -features and <i>trans</i> -factors orchestrating the initial phase of Hox cluster activation	
Chapter 2	Evolutionarily conserved requirement of Cdx for post-occipital tissue emergence	50
Chapter 3	Cdx and T Brachyury co-activate growth signaling in the embryonic axial progenitor niche	72
Chapter 4	Polarized regulatory landscape and Wnt responsiveness underlie Hox activation in embryos	106
Chapter 5	Cdx is a crucial player during colinear Hox activation and defines a trunk segment in the Hox cluster topology	130
Chapter 6	General discussion	152
Addenda		160
	Nederlandse samenvatting	162
	Acknowledgments / Dankwoord	166
	List of publications	172
	Curriculum vitae	175

Summary of the thesis

During vertebrate embryogenesis the body axis extends sequentially from the ‘posterior growth zone’ in which self-renewing axial progenitors reside. Some of these stem cell-like cells are bipotent and contribute to neuroectodermal and mesodermal tissue along the antero-posterior axis of the embryo. In **Chapter 1 (part A)** the transcription factors – among which Cdx, Hox and T Brachyury – and the developmental signals – Wnt, retinoic acid, Fgf – that regulate the maintenance of these ‘neuromesodermal progenitors’ (NMPs) in the posterior growth zone are introduced.

Shortly after the specification of the antero-posterior axis, Hox genes start to be transcribed in the posterior structures of the vertebrate embryo. They are organized in four clusters, and lie in a 3’ to 5’ order in each cluster. From embryonic day (E)7.0 on, the clusters become active on their 3’ side and gradually express their ‘central’ genes to eventually activate the 5’-most Hox13 genes around E9.0. In **Chapter 1 (part B)** we discuss the developmental signals, molecular mechanisms and the *cis*-regulatory domain involved in the onset of the initial Hox gene transcription that engages Hox cluster activation.

Cdx transcription factors are important for tissue generation and axial patterning during embryonic body elongation. In **Chapter 2** we describe mouse embryos in which all three Cdx genes are inactivated. These mutants fail to generate any axial tissue beyond the occipital level as the posterior growth zone loses its ability to maintain the NMP population. The mutant embryos have no problems to activate the 3' Hox genes in the correct spatiotemporal window, but are not able to activate their trunk Hox genes. Additionally, our data indicate that Cdx controls the progenitor niche in the posterior growth zone by regulating Fgf and Wnt signaling pathways.

The Cdx2 and T Brachyury transcription factors are evolutionarily conserved and constitute central players in the genetic network controlling axial growth. They are indispensable for extension of a similar central portion of the axis. In **Chapter 3** we demonstrate that simultaneous loss-of-function of these transcription factors disrupts axial elongation much more severely than each single mutation does. Using embryo-derived epiblast stem cells (EpiSCs), individual and common target loci for Cdx2 and T Brachyury are identified by ChIP-seq and ATAC-seq, and validated *in vivo*.

In **Chapter 4** we demonstrate that the Wnt pathway is the developmental trigger for the transcriptional onset of the 3'-most Hox genes. Upon Wnt stimulation of EpiSCs, temporally colinear expression of the genes of the Hox cluster takes place. Early HoxA activation by Wnt is exerted at the level of the 3' *cis*-regulatory landscape – which is organized in a small topological domain, a 3' subTAD. This subTAD forms the structural basis for multiple layers of 3'-polarized features, including DNA accessibility and enhancer activation, resulting in a 3'-specific Wnt-induced Hox transcription.

In **Chapter 5**, we investigate the role of Cdx genes in the gradual colinear activation of the Hox clusters. The temporally colinear expression of Hox genes is severely affected in EpiSCs derived from Cdx null embryos. We demonstrate that *cis*-regulatory elements around the central Hox genes require Cdx to gain DNA accessibility and to become activated. Several regulatory segments are identified in the Hox neighborhood, forming a structural basis for Hox colinearity of expression. After initial Wnt-driven activation of 3' Hox genes, exerted from a first chromatin segment, Cdx gene products act within a second segment as crucial effectors for the subsequent activation of central Hox genes in the cluster.

Finally, our findings from the preceding chapters are discussed in **Chapter 6**.

Chapter 1

Introduction part A

Region-specific regulation of posterior axial elongation during vertebrate embryogenesis

Roel Neijts¹, Salvatore Simmini¹, Fabrizio Giuliani¹, Carina van Rooijen¹ and Jacqueline Deschamps¹

¹Hubrecht Institute, Developmental Biology and Stem Cell Research, Uppsalalaan 8, 3584 CT Utrecht, and UMC Utrecht

Neijts R, Simmini S, Giuliani F, van Rooijen C and Deschamps J. (2014) Region-specific regulation of posterior axial elongation during vertebrate embryogenesis, *Developmental Dynamics* 243(1):88-98

Abstract

The vertebrate body axis extends sequentially from the posterior tip of the embryo, fueled by the gastrulation process at the primitive streak and its continuation within the tailbud. Anterior structures are generated early, and subsequent nascent tissues emerge from the posterior growth zone and continue to elongate the axis until its completion. The underlying processes have been shown to be disrupted in mouse mutants, some of which were described more than half a century ago. Important progress in elucidating the cellular and genetic events involved in body axis elongation has recently been made on several fronts. Evidence for the residence of self-renewing progenitors, some of which bipotential for neurectoderm and mesoderm, has been obtained by embryo grafting techniques and by clonal analyses in the mouse embryo. Transcription factors of several families including homeodomain proteins have proven instrumental for regulating the axial progenitor niche in the growth zone. A complex genetic network linking these transcription factors and signaling molecules is being unraveled that underlies the phenomenon of tissue lengthening from the axial stem cells. The concomitant events of cell fate decision among descendants of these progenitors begin to be better understood at the levels of molecular genetics and cell behavior. The emerging picture indicates that the ontogenesis of the successive body regions is regulated according to different rules. In addition, parameters controlling vertebrate axial length during evolution have emerged from comparative experimental studies. It is on these issues that this review will focus, mainly addressing the issue of axial extension in the mouse embryo with some comparison with studies in chick and zebrafish, aiming at unveiling the recent progress, and pointing at still unanswered questions for a thorough understanding of the process of embryonic axis elongation.

Progressive anterior to posterior body development in vertebrates

Vertebrates as most bilaterians develop progressively from anterior to posterior. The cephalic structures and the rostral trunk develop first from the anterior epiblast and the mesoderm and endoderm that have arisen early from the primitive streak. At the late primitive streak stage (embryonic day 7.2, E7.2) in the mouse and at definitive streak stage in the chicken (stage 4 according to Hamburger and Hamilton, HH), the streak region will continue contributing descendants for the extending trunk and tail (Schoenwolf, 1977; Kinder et al., 1999). This process corresponds to what has been called the primary body formation. Around closure of the posterior neuropore in both mouse (around E10.0, after some 30 somite pairs have been generated) and chick embryos (at about 14 HH, and 22 somite pairs), tissue emergence will take place from the tailbud rather than from the primitive streak that becomes internalized. The anterior streak/node region subsists as chordo-neural hinge (CNH), and the rest of the streak as ventral ectodermal ridge (VER). Tissue generation for the posteriorly extending lumbosacral region and tail then occurs from the CNH. This phase has been called secondary body formation. Some of the morphogenetic movements during these processes, and gene expression associated with this phase of body elongation suggest that axial extension from the tailbud is the continuation of the earlier process of trunk elongation (Gont et al., 1993; Benazeraf and Pourquie, 2013). In agreement with this idea, mutations affecting the process of axial elongation were shown to affect both the primary and secondary body formation (see below). However, secondary body elongation differs from the primary phase by the fact that the underlying morphological mechanisms rely less on convergence extension and ingression than on the expansion of the three embryological derivatives from progenitors residing in the CNH niche. For instance, the extending neural tube arises by cavitation rather than by lateral elevation and closure of a neural plate. In the following sections we review extensive evidence that the modalities of vertebrate development differ during the laying down of anterior and more posterior tissues, and that there are different rules behind tissue morphogenesis in the successive sub-regions of the axis.

Axial progenitors in the embryonic growth zone

Cell labeling in the node and primitive streak of the chick embryo have long suggested the existence of stem cells at these locations (Selleck and Stern, 1991; Selleck and Stern, 1992). Subsequent experiments following the contribution of single cells in the early somite chick embryo by time-lapse imaging confirmed that stem cell-like progenitors remained in the posterior epiblast whereas their descendants found themselves more anteriorly in the extending neural axis (Mathis et al., 2001). In the mouse, clonal analysis of single epiblast cells in the gastrulating embryo revealed cases of epiblast cells in the anterior primitive streak that contributed descendants along the axis, including some that remained in the node region after a one-day culture period (Lawson et al., 1991; Forlani et al., 2003). Some of these progenitors just posterior to the node were found to give descendants in both neurectoderm and mesoderm (Forlani et al., 2003). Longer term studies by retrospective lineage analysis indi-

cated that some axial progenitors give rise to descendants in differentiated tissues spanning a large rostral-caudal distance and extending back to the node region (Nicolas et al., 1996; Mathis and Nicolas, 2000). The last 10 years have seen major progress in the characterization of axial tissue generation from the primitive streak and tailbud. Serial grafting experiments in early somite mouse and chick embryos demonstrated that stem cell-like progenitors reside in a stem zone located in a small region between the node and the anterior primitive streak (node-streak border, NSB; Fig. 1) and subsequently in the tailbud CNH (Cambray and Wilson, 2002; Cambray and Wilson, 2007; McGrew et al., 2008). These cells contribute descendants for long periods of embryogenesis, to all levels of the elongating trunk in the neural tube and mesoderm. Heterotopic grafts suggested that the properties of these stem cell-like axial progenitors were conferred by the embryonic position rather than being inherent to the cells (Cambray and Wilson, 2007; McGrew et al., 2008). Some cells of the caudo-lateral epiblast (CLE, Fig. 1) indeed will contribute to the stem cell-like axial progenitor population after they have been moved to the NSB by the morphogenetic movement of gastrulation (Cambray and Wilson, 2007). A recent retrospective clonal analysis in the mouse revealed that the only progenitors generating clonal cell populations colonizing most levels of the axis including the stem zone/CNH are bipotent progenitors for both mesoderm and neurectoderm, which reside in the posterior part of the embryo (Tzouanacou et al., 2009). In addition to questioning the dogma concerning the order of derivation of the three germ layers during gastrulation, this study demonstrated that the mode of generation of mesoderm and neural descendants from bipotent progenitors remains the same all along the trunk. These data constitute a solid argument in favor of tailbud development being a continuation of the gastrulation process. Nevertheless, several subclasses of clones descending from single neuro-mesodermal progenitors could be distinguished, revealing a change in composition of the axial progenitor pool active during trunk to tail development (Tzouanacou et al., 2009). Most of the clones contribute descendants to the CNH region where the progenitors reside, but the anterior border of the contribution is variable. The number of clones with an anterior limit of the position of their descendants in the trunk region was higher than that at head and tail levels. This points to an increase in the number of neuro-mesodermal progenitors between gastrulation (when head and neck are laid down) and later stages (early somite stages, when the trunk is generated) (Tzouanacou et al., 2009). It also indicates a decrease in the progenitor pool at tailbud stages when the tail will be formed. The subclasses of clones descending from single neuro-mesodermal progenitors therefore seem to correspond well with the different anatomical regions of the mouse embryo. These data raise the possibility that the composition of the neuro-mesodermal progenitor population changes with time, and thus with the axial level of tissue extension. The rearrangement of the pool of axial stem cells at certain axial levels might correspond to changes in activity of genes essential for modulating these populations. Candidate genes for this modulation are genes the mutation of which impairs axial elongation in a dosage-dependent way, as will be discussed in the following sections. The progenitor pool appears to be depleted in the tailbud at the end of axial extension (Tzouanacou et al., 2009).

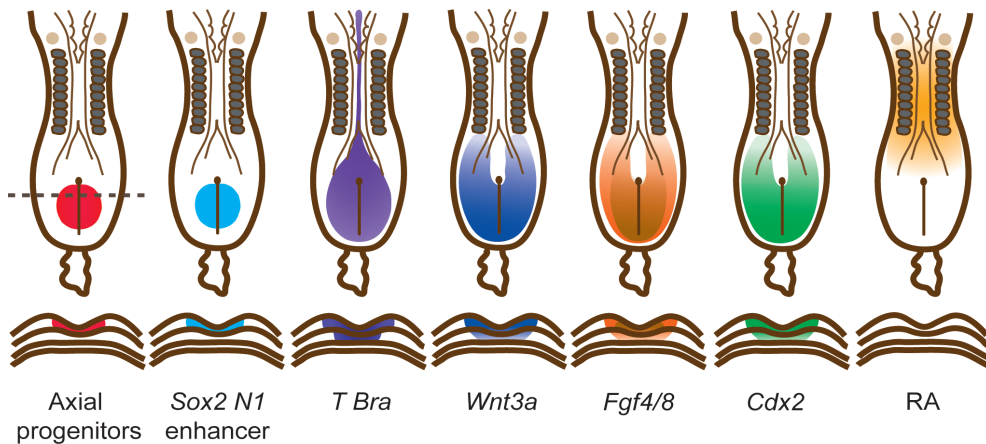


Figure 1: Localization of axial progenitors and expression of molecular players in posterior growth during elongation of the embryonic trunk. Above, schematic renderings of an E8.5 (10-somite) mouse embryo (dorsal view, anterior is up). The primitive streak is represented by a vertical black line, with a black dot at its anterior end, representing the node. From left to right: the red colored area specifies the epiblast cells contributing to the stem cell-like population eventually releasing descendants in mesoderm and neurectoderm of the growing axis; the anterior-most part of the domain (circled) corresponds to the “node-streak border” (NSB), and the rest to the caudo-lateral epiblast (CLE)(Cambray and Wilson, 2007; Wilson et al., 2009); the turquoise blue area corresponds to the activity domain of the *Sox2 N1* enhancer (Takemoto et al., 2011; Kondoh and Takemoto, 2012); purple corresponds to transcription of *T Bra* (Kispert and Herrmann, 1994) and dark blue to the expression of *Wnt3a* (Yoshikawa et al., 1997; Nowotschin et al., 2012); orange with brown sub-domain stands for expression of *Fgf8* (orange) and *Fgf4* (brown), based on (Molotkova et al., 2005; Naiche et al., 2011) and Naiche Adler, Matt Anderson and Mark Lewandoski, personal communication; *Cdx2* expression is shown in green (Young et al., 2009), and the domain of retinoic acid (RA) signaling is shown in ocher, based on *RARβ*-driven reporter gene expression and *Raldh2* activity (Rossant et al., 1991; Niederreither et al., 1997). The lower set of schemes shows the localization of progenitors, gene expression and signals corresponding to the cross section of the upper embryos at the level of the anterior primitive streak (dashed line). The three germ layers are shown, of which the upper layer is epiblast, and the middle layer mesoderm. Below is the primitive endoderm.

Several aspects of axial tissue elongation from the progenitor zone have yet to be elucidated. One of the remaining questions results from the indication that not all tissues at a given axial level derive from the stem cell reservoir localized in the anterior streak/CLE, later the CNH in the tailbud. Tzouanacou and colleagues (Tzouanacou et al., 2009) have demonstrated that this mode of derivation is used by part of the paraxial mesoderm and the neural tube. Regarding the paraxial mesoderm, previous experiments suggested that only the medial parts of the somites arise by a stem cell mode (Selleck and Stern, 1991; Eloy-Trinquet and Nicolas, 2002; Iimura et al., 2007). Cells that will contribute to the lateral part of the somites emerge from a more caudal part of the primitive streak and its flanking epiblast, and are thought to get “organized” subsequently by their medial somitic counterparts when somites are formed (Freitas et al., 2001; Wilson et al., 2009). No evidence for a stem cell-like mode of generation of the lateral plate mesoderm has been found. Homotopic grafting of labeled sub-regions in and along the primitive streak of early somite embryos subsequently cultured for two days (Cambray and Wilson, 2007) confirmed the origin of lateral plate mesoderm at the posterior streak levels established in previous studies (Tam and Beddington, 1987; Kinder et

al., 1999). In embryos grafted with cells of the posterior streak regions, no resident progenitor was left in the primitive streak. This lateral plate mesoderm must therefore emerge from the posterior primitive streak from fully ingressing proximal epiblast. The endoderm also follows its own mode of elongation. Only the early and anterior-most definitive endoderm emerges from the node region (Lawson et al., 1991), whereas the trunk endoderm expands from this early endoderm without stemming from progenitors in the posterior growth zone. Moreover an intercalating contribution of the primitive endoderm to the gut epithelial endoderm has been demonstrated (Kwon et al., 2008). The expansion process of lateral plate mesoderm and endoderm thus differs from that of the stem-cell derived medial mesoderm and neurectoderm. The emergence of all these tissues nevertheless obeys at least some common rules since they are all affected by mutations arresting posterior axial elongation (to be discussed in a later section).

In summary, at least part of the paraxial mesoderm and neurectoderm of the trunk and tail is generated from a population of bipotent progenitors residing in the NSB within the posterior growth zone, with some contribution of the CLE (Fig. 1). The precise population and physical extent of this progenitor zone remains to be defined, and the production mode of trunk lateral mesoderm and definitive endoderm remains to be better understood.

Transcription factors and signals modulating the axial stem cell niche

Spontaneous and induced mutations have long been described (*T Brachyury*, *T Bra* (Dobrovol'skaïa-Zavad'skaïa, 1927); *Danforth Short Tail*, *Sd* (Danforth, 1930; Dunn et al., 1940)) that cause posterior truncation in the mouse embryo. The gene encoding the T-box transcription factor *T Bra* has been identified a long time ago, and it was cloned in the late eighties (Herrmann et al., 1990). The genetic event causing dysgenesis of posterior embryonic tissues in the *Sd* mutant was characterized only very recently, and shown to be a retrotransposon insertion event leading to downregulation of *T Bra* (Lugani et al., 2013; Vlangos et al., 2013). An extensive amount of information has been gathered on the cellular processes and gene expression affected by the *T* gene in the mouse embryo (Beddington et al., 1992; Wilson and Beddington, 1997). *T Bra* is expressed in and along the entire primitive streak and in the notochord (Fig. 1). *T Bra* null epiblast cells are impaired in migrating out of the primitive streak during gastrulation, and their accumulation compromises the process of axial elongation. The defect is dependent on the *T* gene dosage (Wilson and Beddington, 1997). Although *T Bra* is the earliest sign of mesoderm formation in the embryo, *T Bra* null mutants generate anterior mesoderm and form the first 6 somites. This points to a dichotomy between tissue generation at anterior and posterior levels that we will discuss later on. Other experimental investigations made it clear that the Wnt pathway is involved upstream of *T Bra*, in controlling both transcriptional initiation of the gene in the streak epiblast (Rivera-Perez and Magnuson, 2005; Tortelote et al., 2013) and maintenance of its expression (Yamaguchi et al., 1999; Galceran et al., 2001). A break-through discovery was made in studies on the function of T Bra (together with a second *T Bra* gene) in zebrafish. The action of T Bra in embryonic axial elongation was

demonstrated to be mediated by Wnt signaling, meaning that the canonical Wnt pathway acts downstream of *T Bra* upon a process of positive feedback regulation (Martin and Kimelman, 2008). *T Bra* thus acts by maintaining Wnt signaling in the posterior part of the growing embryo. Accordingly, exogenous Wnt signals rescued the posterior truncation of zebrafish *T Bra* mutants (Martin and Kimelman, 2008). These findings are in line with the fact that mouse *Wnt3a* null mutants are as severely posteriorly truncated as *T Bra* null embryos (although some phenotypical features differ between these mutants). A spontaneous hypomorphic mutation of *Wnt3a* (*Vestigial tail, Vt*), believed to disrupt a regulatory element of the *Wnt3a* gene (Greco et al., 1996), also causes posterior axial truncation, albeit in a much lesser extent than the full gene inactivation. Posterior axial growth is therefore dependent, in a dosage dependent way, on the Wnt signaling strength in the embryonic growth zone.

The Wnt pathway is not the only stimulatory effector controlling axial elongation. Mutations inactivating *Fgf receptor 1 (FgfR1)* cause early lethality and truncation of the posterior embryonic regions (Deng et al., 1994; Yamaguchi et al., 1994) among other defects. Hypomorphic mutations in *FgfR1* also lead to posterior truncation of the embryos (Partanen et al., 1998). Active transcription of *Fgf8* in the posterior embryonic tissues takes place exclusively in the primitive streak region. Work in the chicken has shown that transcripts remain present more anteriorly due to their stability (Dubrulle and Pourquie, 2004). During the production of nascent axial tissues extending the axis, a decreasing gradient of *Fgf8* mRNA and protein is thereby generated with a minimum playing the role of wave front specifying a new somite boundary in the presomitic mesoderm (PSM) (Dubrulle and Pourquie, 2004). *Fgf8* transcripts and proteins in chicken and mouse embryos are thus present more widely than where the gene is transcribed (Fig. 1). Conditional inactivation of the *FgfR1* receptor or depletion of both *Fgf4* and *Fgf8* using a Cre recombinase driven by promoters active in the primitive streak also precludes complete axial extension and leads to the downregulation of *T Bra* (Wahl et al., 2007; Naiche et al., 2011; Boulet and Capecchi, 2012). Experiments using chimeric embryos documented that epiblast cells fully deprived of *FgfR1* fail to ingress and instead accumulate in the primitive streak (Ciruna et al., 1997). Later inactivation of Fgfs seems to impair the maintenance of progenitor populations for axial elongation from the tailbud (Boulet and Capecchi, 2012). Wnt and Fgf signaling have both been proposed to act upstream of one another (Aulehla et al., 2003; Wahl et al., 2007), making it difficult to unambiguously build a network involving them together with *T Bra* in axial elongation.

The discovery of the importance of the niche for progenitor maintenance has been a crucial advance in understanding axial elongation. In the mouse this importance was highlighted by the experiments of Wilson and colleagues. They showed that the CLE, a region normally generating mesoderm for a limited extent of the axis, acquires the properties of contributing to both mesoderm and neurectoderm for long periods of time if grafted to the residence of axial stem cells in the anterior-most part of the streak (Cambray and Wilson, 2007). The lack of molecular definition of the niche of axial stem cells, the very specific area at the anterior extremity of the primitive streak supporting the maintenance of the long-term axial

progenitors (NSB and CNH) has puzzled the field for years. In the absence of unique markers, attempts to identify genetic features for this zone have kept the attention on important players in axial extension such as *Wnt3a* and *Fgf8*, and on the co-expression of genes associated with early pluripotency and with early mesoderm differentiation such as *Sox2* and *T Bra*, respectively, in the mouse (Cambray and Wilson, 2007; Wilson et al., 2009) and in the chicken embryo (Delfino-Machin et al., 2005; Olivera-Martinez et al., 2012). However, all these players are each expressed more broadly than the territory defined for the axial stem cell niche (Fig. 1). The transcriptional activity of an early enhancer (*N1*) of the *Sox2* gene seems to overlap with the anterior streak and CLE, and is therefore also not restricted to the long-term axial progenitor zone (Takemoto et al., 2011)(Fig. 1). The identification of long-term axial progenitors has remained elusive, as it has not been possible so far to isolate self-renewing axial stem cells and define their properties and transcriptome. Co-expression studies with *Sox2* and *T Bra* by immuno-histochemistry are so far the most likely to define a sub-population of cells that can be considered as multipotent axial stem cells, as has been recently shown in the chick tailbud by Olivera Martinez et al. (Olivera-Martinez et al., 2012).

Cell fate decision and germ layer expansion during axial extension

Cell fate choice and appropriate differentiation are as crucial for axial growth as progenitor self-renewal. The evidence for long-term neuro-mesodermal progenitors residing in the anterior streak/CLE (later the CNH) (Cambray and Wilson, 2002; Cambray and Wilson, 2007; McGrew et al., 2008), and the recent discovery of the close kinship between neurectoderm and mesoderm among the derivatives of these axial progenitors (Tzouanacou et al., 2009; Kondoh and Takemoto, 2012; Martin and Kimelman, 2012) were key to understanding the control of cell fate choice during axial extension. All mutants mentioned in the previous sections, which compromise posterior axial elongation exhibit ectopic neurectoderm tissues at locations where mesoderm normally ends up. *T Bra* mutant embryos (Yamaguchi et al., 1999), *Wnt3a* mutants (Yoshikawa et al., 1997), and chimeric embryos containing *Fgfr1* mutant epiblast cells (Ciruna et al., 1997) all formed ectopic neural tubes at their posterior-most axial levels. The molecular genetic mechanism underlying this deviation from the normal behavior, and the interdependence between mesoderm and neurectoderm differentiation has been recently elucidated (Takemoto et al., 2011; Kondoh and Takemoto, 2012). The early *N1* enhancer of the *Sox2* gene, active in the axial stem cells of the CLE, is normally turned off as soon as cells ingress through the primitive streak (Takemoto et al., 2011) under the repressive influence of *Tbx6* in the mesoderm. If *Tbx6* is absent, *N1* remains active upon cell ingression through the streak, and neural structures develop instead of PSM and somites. In that case, axial extension fails because of an arrest of paraxial mesoderm production. Work in zebrafish showed that canonical Wnt signaling is required for mesodermal differentiation of the bipotent neuro-mesoderm axial progenitors (Martin and Kimelman, 2012). Inactivating Wnt signaling in these progenitors causes the descendants to form neural tissue exclusively and to remain in the epiblast instead of undergoing epithelium-mesenchymal transition (EMT).

Recent work using genetic analysis of single and double mouse mutants proposed ordering in some at least of these events involved in axial elongation. *Wnt3a* was suggested to regulate the process of ingression by EMT, and *Tbx6* was proposed to act at a hierarchically lower level in post-ingression mesoderm (Nowotschin et al., 2012). Maintenance of the axial stem cell population by the Wnt pathway is also a functional component of axial growth, which would not distinguish itself phenotypically very much from a function in EMT. Arrest of epiblast ingression through the primitive streak and exhaustion of the axial stem progenitor pool at the streak level are therefore expected to be difficult to tell apart.

In summary, progress has been made in understanding cell fate decision during derivation of posterior neural and mesodermal tissues from the axial stem cells. It has become clear that the Wnt and Fgf pathways modulate the mesoderm/neurectoderm differentiation potential of the long-term progenitors. This modulation affects the balance between tissue production and differentiation regulating the extension of the body axis.

Cdx and Hox genes in axial elongation

The involvement of additional transcription factors in steering axial extension via stimulation of Wnt and Fgf signaling recently added a level of complexity in the underlying genetic network. Cdx genes (orthologs of *Drosophila caudal*) are expressed in the posterior growth zone similarly to *Wnt3a* and *Fgf8* (Young et al., 2009) (Fig. 1). Mutations inactivating the three Cdx genes impair axial extension with a severity that depends on the combination of alleles inactivated (see Chapter 2 of this thesis) (Savory et al., 2009; Young et al., 2009; Savory et al., 2011; van de Ven et al., 2011; van Rooijen et al., 2012). Failure to complete the elongation of posterior tissues in Cdx mutants is at least partially rescued by a genetic gain of function of the Wnt pathway, or by exogenous Fgf added to the mutant embryos in culture (Young et al., 2009; van Rooijen et al., 2012). The realm of action of the Cdx genes encompasses the mouse embryonic trunk and tail extent of the axis. Again, this points to the difference, mentioned in an earlier section, in the genetic mechanism underlying the generation of the posterior axial tissues relatively to the more anterior cephalic and occipital structures. The loss of all three Cdx genes is accompanied by extinction of *T Bra* expression after 5 somites have been generated (Chapter 2). Although Cdx binding sequences were identified in the proximal upstream region of the *T Bra* gene (Savory et al., 2009), *T Bra* is normally initiated, and remains expressed until early somite stages in Cdx null mutants (Chapter 2). These data might indicate that Cdx proteins work on maintaining *T Bra* transcription in the posterior growth zone during trunk development, but are dispensable for the initial expression of the *T* gene in the primitive streak at the beginning of gastrulation. Alternatively, the loss of Cdx might cause the exhaustion of cells with progenitor properties in the streak region, causing the disappearance of the tissue that expresses *T Bra* in the primitive streak. The expression of *T Bra* in the notochord is not affected in Cdx null embryos (Chapter 2). Like mutations in *T Bra*, and impairment of the Wnt and Fgf pathways, Cdx deficiencies affect the neural/mesoderm differentiation choice, since partial Cdx mutants exhibit ectopic neural tissues (van de Ven et

al., 2011). The data collected on the Cdx mutants therefore ascertain that Cdx proteins affect the activity of the self-renewing bipotent axial progenitors, acting on their maintenance and on their differentiation choice during axial growth. T Bra, Cdx, Wnt and Fgf all are needed in a dosage-dependent way to promote posterior axial elongation. These gene products thus possess the properties expected from regulators fine tuning the pool of neuro-mesodermal axial progenitors and their differentiation, as proposed by Tzouanacou et al. (Tzouanacou et al., 2009).

Essential questions remain to be answered in order to understand the precise molecular interactions between the Cdx transcription factors and the key effector Wnt and Fgf pathways. Reporter assays with *Wnt3a* promoter fragments in transfected cells in culture suggested a direct role of Cdx in stimulating *Wnt3a* transcription (Savory et al., 2009). However, no specific impact of Cdx mutations on components of the Wnt and Fgf pathways has been demonstrated so far, in spite of several transcriptome analyses in posterior tissues of Cdx mutant embryos (van de Ven et al., 2011).

Another complication concerns the functional relationship between Cdx and their relatives the Hox genes (Young et al., 2009). Hox genes are also expressed in the posterior growth zone during the laying down of the trunk and tail (Young et al., 2009). It was shown that Hox genes of the middle of the clusters (*Hoxb8* for instance) can rescue the posterior truncation of Cdx mutants. It is not known so far whether these Hox genes and Cdx activate the same or overlapping series of downstream targets genes. Striking in any case is the fact that the last genes of the Hox clusters exert an opposite influence on axial growth. A precocious activation of these Hox13 genes is not tolerated in the posterior growth zone (Young et al., 2009). Hox13 genes exert a dominant negative effect on the activity of Cdx and on earlier Hox cluster members, antagonizing their function. This situation fits the principle of temporal collinearity of action of the Hox genes (Duboule, 1994; Iimura and Pourquie, 2006; Duboule, 2007; Tschopp et al., 2009). Hox13 genes are candidates to normally initiate the slowing down of axial extension at the trunk/tail transition, foreshadowing the arrest of posterior growth. In the genetic network underlying the control of axial elongation, the Hox/Cdx genes would bring in the notion of timing and positional identity of the axial tissues that are being generated. Cdx and Hox genes would integrate the spatial and temporal components that control axial growth.

The elucidation of the role of Cdx and Hox genes in the interacting network of players at work during axial growth and tissue differentiation will have to await a more complete molecular genetic analysis of the available single and multiple mouse mutants. It will also need a more exhaustive mapping of the molecular interactions between the different transcription factors and their genomic targets.

Control of body length during evolution

The experiments of Cambray and Wilson (Cambray and Wilson, 2002) have shown that the long-term neuro-mesodermal progenitors start losing their activity at the trunk/tail transi-

tion, upon a genetically programmed decrease of Wnt and Fgf signaling in the “aging” progenitor niche in growing mouse embryos. This transition corresponds to the shrinking of the PSM, which was shown in the chick to anticipate the termination of somitogenesis and posterior tissue addition (Gomez et al., 2008). This transition also correlates with a milestone in the activation of clustered Hox genes: the activation of Hox13 genes. These genes are believed to promote the termination of axial growth (Young et al., 2009). Comparative analysis of Hox gene expression during embryogenesis of the corn snake and shorter body animals has revealed that two of the four Hox13 paralogs are not expressed in the growth zone in snake embryos, whereas all four are expressed in the related shorter body squamate, the whiptail lizard (Di-Poi et al., 2010). These findings suggest that loss of expression of two of the Hox13 genes might correlate with a delay in unleashing the slowing down of axial extension, which would result in an extended trunk elongation. The presence of a high number of DNA repeats within the snake Hox clusters, which are not found in the mammalian clusters (Di-Poi et al., 2010) suggest that relaxing the tightly controlled temporal progression of Hox gene expression has played a role in the emerging of long body vertebrates during evolution.

The role of the Hox-related Cdx genes in body axis growth seems to have been evolutionary conserved. The requirement of *caudal*/Cdx for generating axial tissues posterior to the head was not only demonstrated in mice, but also in short (the flour beetle) and intermediate (the cricket) germ band insects, and in the crustacean *Artemia* (Copf et al., 2004; Shinmyo et al., 2005). The ancestors of these phylogenetically distant species thus already involved Cdx transcription factors to construct their trunk and posterior structures. This demonstrates that axial extension from a posterior growth zone is the basal mode of body axis elongation in bilateria (Akam, 1989; Jacobs et al., 2005) and that Cdx/Wnt/Fgf constitutes an evolutionary conserved genetic toolkit for trunk and tail generation (Copf et al., 2004; Shinmyo et al., 2005; Beermann and Schroder, 2008; Beermann et al., 2011).

Role of retinoids in the Wnt/Fgf control of axial lengthening

A signaling route not involved so far in this overview is that of retinoic acid (RA) and its pathway components. Tight regulation of this pathway is known to be essential for well controlled axial growth, and to allow interaction of the pathway with every single effector mentioned above. The transcription factor-encoding genes *T Bra* and Cdx, and Wnt and Fgf signaling all were shown to be affected by loss or gain of RA biosynthesis (Abu-Abed et al., 2001; Sakai et al., 2001; Diez del Corral and Storey, 2004; Vermot et al., 2005; Ribes et al., 2009; Zhao and Duester, 2009). RA signaling in the early embryo is spatially restricted by RA synthesizing and metabolizing enzymes, *Raldh2* and *Cyp26a1* respectively. RA-responsive domains are highly dynamic during the time course of posterior development from primitive streak stages and somitogenesis to axis termination.

At early somite stages, complementary domains of RA biosynthesis and degradation are generated in the posterior part of the embryo. Somites and anterior PSM express *Raldh2* that generates RA, which then diffuses in the overlying neural plate (Sirbu and Duester, 2006;

Ribes et al., 2009). These activities of source and sink create a gradient of RA concentrations with a maximum in the last somite, fading out posteriorly in the primitive streak and stem zone region, where *Cyp26a1* is expressed. The RA distribution is thus opposite to the *Fgf8* gradient mentioned in an earlier section of this review. During trunk formation, the dichotomy between a posterior stem cell zone expressing Fgf, Wnt, T Bra, Cdx and *Cyp26a1*, and a differentiation zone with high RA bioactivity, is therefore very clear. Cells leaving the posterior progenitor zone become exposed to higher RA and lower *Fgf8* levels. They then undergo differentiation into the caudal neural plate, or segment into somites. RA is cleared in the stem cell niche by *Cyp26a1*, the expression of which depends on Fgf signaling (Wahl et al., 2007; Boulet and Capecchi, 2012) and on Cdx (Savory et al., 2009; Young et al., 2009; van Rooijen et al., 2012) and T Bra (Martin and Kimelman, 2010; Vidigal et al., 2010). Exposure of the embryos to supra-physiological levels of RA by maternal administration, or following a defective RA metabolism, causes severe axial truncations (Kessel and Gruss, 1991; Shum et al., 1999; Abu-Abed et al., 2001; Sakai et al., 2001). Transient RA treatment during somitogenesis restricts Fgf and Wnt signaling and expression of *T Bra* in the progenitor zone (Iulianella et al., 1999; Shum et al., 1999; Diez del Corral et al., 2003; Wilson et al., 2009). Reciprocally, Fgf antagonizes RA signaling, at least in part by repressing *Raldh2* expression in the region where these players overlap (Diez del Corral et al., 2003; Wahl et al., 2007; Olivera-Martinez et al., 2012). Thus, a large amount of experimental evidence shows that RA is necessary to induce neural differentiation from the axial progenitor descendants, and that RA clearance in the growth zone is required to maintain the genetic program of progenitors during trunk formation.

At the trunk/tail transition (E10), Fgf and Wnt signals decline, the PSM shrinks (Gomez et al., 2008), and the RA-producing domain then abuts the progenitor zone in the tailbud. *Cyp26a1* levels decrease posteriorly at this stage, in both chicken and mouse embryos. Surprisingly, *Raldh2* was found to become transiently induced in a localized area of the tailbud (Tenin et al., 2010; Olivera-Martinez et al., 2012). RA at this location was reported to be biologically active in the chick, as it leads to a decline in the expression of Fgf, Wnt and *T Bra*, apoptosis and axis termination (Olivera-Martinez et al., 2012). In co-culture experiments chick tailbuds have been shown to activate RARE-*lacZ* in a reporter cell line, whereas mouse tailbuds did not appear to be an endogenous source of RA in the same assay (Tenin et al., 2010). It might be that RA does not play a role in the termination of the mouse body axis. In *Raldh2* mutant embryos transiently treated with RA to overcome early developmental lethality, posterior *Fgf8* is downregulated and axial elongation is arrested normally (Cunningham et al., 2011). The mechanism of axial termination is thus possibly different in these two organisms. It is not known how the differences in the mechanism of axial termination have arisen during evolution. Altogether the data again emphasize the differential regulation of the distinct anatomical regions of the body axis.

Conclusions and additional future prospects

Trunk and tail tissues in vertebrates stem from the embryonic posterior growth zone that harbors a niche for the maintenance of axial progenitors, some of which are self-renewing and bipotent for mesoderm and neurectoderm. The progenitor niche ages at mid-gestation, upon genetically programmed decrease of Wnt and Fgf signaling. These processes foreshadow the termination of axial growth. The aging process can be reversed, as the niche and its progenitors can be rejuvenated upon transplantation into the corresponding region of a younger embryo (Cambray and Wilson, 2002). The molecular genetic characterization of the long-term progenitor niche, within the region of overlap between the activity domains of the key effectors of axial growth, is an important remaining goal in future research. Given the fact that the effectors in the progenitor niche, Wnt and Fgf, are regulated by the transcription factors *T Bra* and *Cdx/Hox* in feed-back loop relationships (Aulehla et al., 2003; Wahl et al., 2007), the transcriptional control of these transcription factors in the growth zone is a major steering component in axial growth. Understanding the regulation of these regulators might shed light on the upstream command of key events of posterior morphogenesis such as the emergence of the *Fgf8* gradient that is crucial for posterior axial growth (Diez del Corral and Storey, 2004; Dubrulle and Pourquie, 2004; Naiche et al., 2011; Boulet and Capecchi, 2012). A pioneer study regarding the transcriptional regulation of the *Fgf8* locus has just been published (Marinic et al., 2013) and promises to lead to unveiling regulatory events driving posterior embryonic growth. The establishment of the epistatic relationship between the players in the network driving axial extension, and the precise mapping of the productive binding events of the key transcription regulators at their genomic targets will be essential to fully understand the genetic program underlying axial tissue growth and patterning.

A conclusion of the data reviewed above is that the rules governing axial elongation anterior to the trunk, within the trunk and at the termination level of the tail follow different principles. The phenotype of *T Bra* and *Cdx* null mutants indicates that the corresponding gene products are needed for axial elongation posterior to the 5/6 most anterior somites. This suggests a scenario whereby early Wnt signaling (*Wnt3*) is enough to ensure extension of the anterior axial structures, whereas trunk and posterior elongation requires prolonged Wnt signaling (*Wnt3a*) maintained by *T Bra* and *Cdx*. Early lethality of the *Wnt3* mutants that fail to undergo gastrulation (Liu et al., 1999) precludes assessing the validity of this hypothesis. The same may hold true for Fgf signaling. Early *Fgf4* and *Fgf8* may suffice for occipital tissue elongation but fall short of supporting more posterior trunk and tail growth when they are inactivated shortly thereafter by *T Bra*-driven Cre recombination (Naiche et al., 2011). Fgf and Wnt may therefore play a continued role in maintaining the progenitors for axial extension in the different windows of developmental time. Boosting of both pathways by *Cdx* and *T Bra* might mark the occipital/post-occipital transition. An unknown in this scenario is whether anterior tissues are laid down at all from a population of long-term neuro-mesodermal precursors.

Tissue morphogenesis in body regions at different axial levels has been shown ear-

lier to obey to distinct rules, and to differentially depend on specific unravelled molecular genetic circuits. The oscillatory activity of genes inherent to the segmentation clock such as *Lunatic fringe* is differentially required for thoraco-lumbar and sacro-caudal body patterning of the presomitic mesoderm (Shifley et al., 2008; Stauber et al., 2009). Shifley and colleagues suggest that the process of primary versus secondary body formation may impact on the regulation of rostro-caudal patterning by the segmentation clock during the various stages of anteroposterior axis development. Not only the sclerotome, but also the myotome seems to be differentially controlled during primary and secondary body formation: *PDGF α* receptor mutants were found to exhibit myotome defects in the 20 most rostral somites and to appear normal in the caudal somites (Soriano, 1997). Myogenesis has indeed been shown to be controlled by different strategies during ontogenesis of the different regions of body development. Trunk and head muscles are generated by paraxial mesoderm that is respectively segmented and unsegmented (at least not segmented in the same way as the trunk) into somites (reviewed by (Sambasivan et al., 2011)). Cranial and somitic myogenesis clearly depends on distinct genetic networks (Sambasivan et al., 2011). Comparison of these processes in different chordates suggests that the neck muscles appeared at a transition zone between the cranial and trunk mesoderm and ancestral vertebrates do not seem to have possessed neck muscles (Sambasivan et al., 2011). Strikingly the endoderm of the developing embryonic body has been recently shown also to differentially depend on well-defined signaling pathways. Wnt/ β -catenin signaling is crucially important for the formation of definitive endoderm of the mid- and hindgut, whereas it is dispensable for foregut formation (Engert et al., 2013).

Another difference between the progenitors of the anterior-most, and trunk part of the axis concerns their sensitivity to RA. The early axial progenitor niche at primitive streak stages expresses *Raldh2* (Vermot et al., 2005; Ribes et al., 2009) whereas the presence of RA in the growth zone during trunk elongation is detrimental to the progenitors (see above). The biphasic response of embryonic axial progenitors exhibiting initial compatibility of RA exposure with self-renewal, and later drifting towards differentiation in the presence of RA was reported to be observed as well in murine embryonic stem cells in culture (Stavridis et al., 2010). These dynamic changes in niche composition with developmental time make it even more difficult than appreciated so far to understand the essential properties of these axial progenitors and their niche.

Many aspects of tissue generation from the progenitor zone remain to be elucidated, as set out in the different sections above. In addition, the recent arousal of interest for additional parameters to be considered in the control of morphogenesis is expected to bring new light on the phenomenon of posterior embryonic growth. In addition to the involvement of cell adhesion, known to be controlled during gastrulation, cell flow and fluidity change in posterior tissues during axial elongation have recently been evoked as potential regulators of axial growth in zebrafish embryos (Lawton et al., 2013). The authors found that the movement of cells from the dorso-medial zone – the remnant site of gastrulation – to the PSM depends on decreasing Fgf and Wnt signal concentrations. This migration within the tailbud is accom-

panied by a change in cell flow and tissue fluidity. Previous work in chicken embryos showed that high Fgf signaling promotes cell motility in the posterior PSM, that this motility decreases as cells approach the segmentation boundary in the low values of the Fgf gradient and that disruption of the motility gradient results in slowing down of axial elongation (Benazeraf et al., 2010). The distribution of non-canonical Wnt and Fgf signaling, known to affect cell migration directly or indirectly, would generate a balanced equilibrium between the cell flow rate, the emergence of differentiated derivatives and the coherence of collective cell migration within the tailbud. Disrupting the Wnt and Fgf gradients would introduce chaos in cell flux, aberrant elongation of the trunk, and in some cases inappropriate localization of tissue derivatives such as neurectoderm. It is evident that we are far from completely apprehending the complexity of signaling regulating the morphogenetic event of posterior axial growth. Future progress in these respects is heavily awaited. The design of non-intrusive methods to genetically label selected groups of cells in developing embryos in culture, and the development of live imaging technology during progressing embryogenesis *in vivo* should allow to precise the role of each effector in the process of axial elongation.

Acknowledgments

We wish to thank the anonymous reviewers for their helpful comments on the manuscript. We also thank Naiche L. Adler, Matthew J. Anderson and Mark Lewandoski (National Cancer Institute, Frederick, Maryland, USA) for kindly sharing unpublished data concerning *Fgf4* and *Fgf8* expression. The authors work at the Hubrecht Institute and Utrecht University Medical Center, and benefit from grants from NWO/ALW and NIRM (Netherlands Institute for Regenerative Medicine, grant No. FES0908).

References

- Abu-Abed S, Dolle 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.
- Akam M. 1989. Hox and HOM: homologous gene clusters in insects and vertebrates. *Cell* 57:347-349.
- Aulehla A, Wehrle C, Brand-Saberi B, Kemler R, Gossler A, Kanzler B, Herrmann BG. 2003. Wnt3a plays a major role in the segmentation clock controlling somitogenesis. *Dev Cell* 4:395-406.
- Beddington RS, Rashbass P, Wilson V. 1992. Brachyury - a gene affecting mouse gastrulation and early organogenesis. *Dev Suppl*:157-165.
- Beermann A, Pruhs R, Lutz R, Schroder R. 2011. A context-dependent combination of Wnt receptors controls axis elongation and leg development in a short germ insect. *Development* 138:2793-2805.
- Beermann A, Schroder R. 2008. Sites of Fgf signalling and perception during embryogenesis of the beetle *Tribolium castaneum*. *Dev Genes Evol* 218:153-167.
- Benazeraf B, Francois P, Baker RE, Denans N, Little CD, Pourquie O. 2010. A random cell motility gradient downstream of FGF controls elongation of an amniote embryo. *Nature* 466:248-252.
- Benazeraf B, Pourquie O. 2013. Formation and Segmentation of the Vertebrate Body Axis. *Annu Rev Cell Dev Biol*.
- Boulet AM, Capecchi MR. 2012. Signaling by FGF4 and FGF8 is required for axial elongation of the mouse embryo. *Dev Biol* 371:235-245.
- Cambray N, Wilson V. 2002. Axial progenitors with extensive potency are localised to the mouse chordoneural hinge. *Development* 129:4855-4866.
- Cambray N, Wilson V. 2007. Two distinct sources for a population of maturing axial progenitors. *Development* 134:2829-2840.
- Ciruna BG, Schwartz L, Harpal K, Yamaguchi TP, Rossant J. 1997. Chimeric analysis of fibroblast growth factor receptor-1 (Fgfr1) function: a role for FGFR1 in morphogenetic movement through the primitive streak. *Development* 124:2829-2841.
- Copf T, Schroder R, Averof M. 2004. Ancestral role of caudal genes in axis elongation and segmentation. *Proc Natl Acad Sci U S A* 101:17711-17715.
- Cunningham TJ, Zhao X, Duester G. 2011. Uncoupling of retinoic acid signaling from tailbud development before termination of body axis extension. *Genesis* 49:776-783.
- Danforth CH. 1930. Developmental anomalies in a special strain of mice. *American Journal of Anatomy* 45:275-287.
- Delfino-Machin M, Lunn JS, Breitzkreuz DN, Akai J, Storey KG. 2005. Specification and maintenance of the spinal cord stem zone. *Development* 132:4273-4283.
- Deng CX, Wynshaw-Boris A, Shen MM, Daugherty C, Ornitz DM, Leder P. 1994. Murine FGFR-1 is required for early postimplantation growth and axial organization. *Genes Dev* 8:3045-3057.
- Di-Poi N, Montoya-Burgos JI, Miller H, Pourquie O, Milinkovitch MC, Duboule D. 2010. Changes in Hox genes' structure and function during the evolution of the squamate body plan. *Nature* 464:99-103.
- Diez del Corral R, 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.
- Diez del Corral R, Storey KG. 2004. Opposing FGF and retinoid pathways: a signalling switch that controls differentiation and patterning onset in the extending vertebrate body axis. *Bioessays* 26:857-869.
- Dobrovolskaia-Zavadskaia N. 1927. Sur la mortification spontanee de la chez la souris nouveau-nee et sur l'existence d'un caractere (facteur) hereditaire, non-viable. *Crit Rev Soc Biol* 97:114-116.
- Duboule D. 1994. Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. *Dev Suppl*:135-142.
- Duboule D. 2007. The rise and fall of Hox gene clusters. *Development* 134:2549-2560.
- Dubrulle J, Pourquie O. 2004. fgf8 mRNA decay establishes a gradient that couples axial elongation to patterning in

- the vertebrate embryo. *Nature* 427:419-422.
- Dunn LC, Schofnheimer SG, Bryson V. 1940. A new mutation in the mouse affecting spinal column and urogenital system. *The Journal of Heredity* 31:343-348.
- Eloy-Trinquet S, Nicolas JF. 2002. Cell coherence during production of the presomitic mesoderm and somitogenesis in the mouse embryo. *Development* 129:3609-3619.
- Engert S, Burtcher I, Liao WP, Dulev S, Schotta G, Lickert H. 2013. Wnt/beta-catenin signalling regulates Sox17 expression and is essential for organizer and endoderm formation in the mouse. *Development*.
- Forlani S, Lawson KA, Deschamps J. 2003. Acquisition of Hox codes during gastrulation and axial elongation in the mouse embryo. *Development* 130:3807-3819.
- Freitas C, Rodrigues S, ChARRIER JB, Teillet MA, Palmeirim I. 2001. Evidence for medial/lateral specification and positional information within the presomitic mesoderm. *Development* 128:5139-5147.
- Galceran J, Hsu SC, Grosschedl R. 2001. Rescue of a Wnt mutation by an activated form of LEF-1: regulation of maintenance but not initiation of Brachyury expression. *Proc Natl Acad Sci U S A* 98:8668-8673.
- Gomez C, Ozbudak EM, Wunderlich J, Baumann D, Lewis J, Pourquie O. 2008. Control of segment number in vertebrate embryos. *Nature* 454:335-339.
- Gont LK, Steinbeisser H, Blumberg B, de Robertis EM. 1993. Tail formation as a continuation of gastrulation: the multiple cell populations of the *Xenopus* tailbud derive from the late blastopore lip. *Development* 119:991-1004.
- Greco TL, Takada S, Newhouse MM, McMahon JA, McMahon AP, Camper SA. 1996. Analysis of the vestigial tail mutation demonstrates that Wnt-3a gene dosage regulates mouse axial development. *Genes Dev* 10:313-324.
- Herrmann BG, Labeit S, Poustka A, King TR, Lehrach H. 1990. Cloning of the T gene required in mesoderm formation in the mouse. *Nature* 343:617-622.
- Imura T, Pourquie O. 2006. Collinear activation of Hoxb genes during gastrulation is linked to mesoderm cell ingression. *Nature* 442:568-571.
- Imura T, Yang X, Weijer CJ, Pourquie O. 2007. Dual mode of paraxial mesoderm formation during chick gastrulation. *Proc Natl Acad Sci U S A* 104:2744-2749.
- Iulianella A, Beckett B, Petkovich M, Lohnes D. 1999. A molecular basis for retinoic acid-induced axial truncation. *Dev Biol* 205:33-48.
- Jacobs DK, Hughes NC, Fitz-Gibbon ST, Winchell CJ. 2005. Terminal addition, the Cambrian radiation and the Phanerozoic evolution of bilaterian form. *Evol Dev* 7:498-514.
- Kessel M, Gruss P. 1991. Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. *Cell* 67:89-104.
- Kinder SJ, Tsang TE, Quinlan GA, Hadjantonakis AK, Nagy A, Tam PP. 1999. The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo. *Development* 126:4691-4701.
- Kispert A, Herrmann BG. 1994. Immunohistochemical analysis of the Brachyury protein in wild-type and mutant mouse embryos. *Dev Biol* 161:179-193.
- Kondoh H, Takemoto T. 2012. Axial stem cells deriving both posterior neural and mesodermal tissues during gastrulation. *Curr Opin Genet Dev* 22:374-380.
- Kwon GS, Viotti M, Hadjantonakis AK. 2008. The endoderm of the mouse embryo arises by dynamic widespread intercalation of embryonic and extraembryonic lineages. *Dev Cell* 15:509-520.
- Lawson KA, Meneses JJ, Pedersen RA. 1991. Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development* 113:891-911.
- Lawton AK, Nandi A, Stulberg MJ, Dray N, Sneddon MW, Pontius W, Emonet T, Holley SA. 2013. Regulated tissue fluidity steers zebrafish body elongation. *Development* 140:573-582.
- Liu P, Wakamiya M, Shea MJ, Albrecht U, Behringer RR, Bradley A. 1999. Requirement for Wnt3 in vertebrate axis formation. *Nat Genet* 22:361-365.
- Lugani F, Arora R, Papeta N, Patel A, Zheng Z, Sterken R, Singer RA, Caridi G, Mendelsohn C, Sussel L, Papaioannou

- VE, Gharavi AG. 2013. A retrotransposon insertion in the 5' regulatory domain of *Ptf1a* results in ectopic gene expression and multiple congenital defects in Danforth's short tail mouse. *PLoS Genet* 9:e1003206.
- Marinic M, Aktas T, Ruf S, Spitz F. 2013. An integrated holo-enhancer unit defines tissue and gene specificity of the *fgf8* regulatory landscape. *Dev Cell* 24:530-542.
- Martin BL, Kimelman D. 2008. Regulation of canonical Wnt signaling by Brachyury is essential for posterior mesoderm formation. *Dev Cell* 15:121-133.
- Martin BL, Kimelman D. 2010. Brachyury establishes the embryonic mesodermal progenitor niche. *Genes Dev* 24:2778-2783.
- Martin BL, Kimelman D. 2012. Canonical Wnt signaling dynamically controls multiple stem cell fate decisions during vertebrate body formation. *Dev Cell* 22:223-232.
- Mathis L, Kulesa PM, Fraser SE. 2001. FGF receptor signalling is required to maintain neural progenitors during Hensen's node progression. *Nat Cell Biol* 3:559-566.
- Mathis L, Nicolas JF. 2000. Different clonal dispersion in the rostral and caudal mouse central nervous system. *Development* 127:1277-1290.
- McGrew MJ, Sherman A, Lillico SG, Ellard FM, Radcliffe PA, Gilhooley HJ, Mitrophanous KA, Cambay N, Wilson V, Sang H. 2008. Localised axial progenitor cell populations in the avian tail bud are not committed to a posterior Hox identity. *Development* 135:2289-2299.
- Molotkova N, Molotkov A, Sirbu IO, Duester G. 2005. Requirement of mesodermal retinoic acid generated by *Raldh2* for posterior neural transformation. *Mech Dev* 122:145-155.
- Naiche LA, 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.
- Nicolas JF, Mathis L, Bonnerot C, Saurin W. 1996. Evidence in the mouse for self-renewing stem cells in the formation of a segmented longitudinal structure, the myotome. *Development* 122:2933-2946.
- Niederreither K, McCaffery P, Drager UC, Chambon P, Dolle P. 1997. Restricted expression and retinoic acid-induced downregulation of the retinaldehyde dehydrogenase type 2 (RALDH-2) gene during mouse development. *Mech Dev* 62:67-78.
- Nowotschin S, Ferrer-Vaquero A, Concepcion D, Papaioannou VE, Hadjantonakis AK. 2012. Interaction of *Wnt3a*, *Msn1* and *Tbx6* in neural versus paraxial mesoderm lineage commitment and paraxial mesoderm differentiation in the mouse embryo. *Dev Biol* 367:1-14.
- Olivera-Martinez I, Harada H, Halley PA, Storey KG. 2012. Loss of FGF-dependent mesoderm identity and rise of endogenous retinoid signalling determine cessation of body axis elongation. *PLoS Biol* 10:e1001415.
- 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. *Genes Dev* 12:2332-2344.
- Ribes V, Le Roux I, Rhinn M, Schuhbauer B, Dolle P. 2009. Early mouse caudal development relies on crosstalk between retinoic acid, *Shh* and *Fgf* signalling pathways. *Development* 136:665-676.
- Rivera-Perez JA, Magnuson T. 2005. Primitive streak formation in mice is preceded by localized activation of Brachyury and *Wnt3*. *Dev Biol* 288:363-371.
- Rossant J, Zirngibl R, Cado D, Shago M, Giguere V. 1991. Expression of a retinoic acid response element-hsplaZ transgene defines specific domains of transcriptional activity during mouse embryogenesis. *Genes Dev* 5:1333-1344.
- 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 antero-posterior axis within the mouse embryo. *Genes Dev* 15:213-225.
- Sambasivan R, Kuratani S, Tajbakhsh S. 2011. An eye on the head: the development and evolution of craniofacial muscles. *Development* 138:2401-2415.
- Savory JG, Bouchard N, Pierre V, Rijli FM, De Repentigny Y, Kothary R, Lohnes D. 2009. *Cdx2* regulation of posterior development through non-Hox targets. *Development* 136:4099-4110.

- Savory JG, Mansfield M, Rijli FM, Lohnes D. 2011. Cdx mediates neural tube closure through transcriptional regulation of the planar cell polarity gene Ptk7. *Development* 138:1361-1370.
- Schoenwolf GC. 1977. Tail (end) bud contributions to posterior region of chick-embryo. *Anatomical Record* 187:708-708.
- Selleck MA, Stern CD. 1991. Fate mapping and cell lineage analysis of Hensen's node in the chick embryo. *Development* 112:615-626.
- Selleck MAJ, Stern CD. 1992. Commitment of mesoderm cells in Hensens node of the chick-embryo to notochord and somite. *Development* 114:403-415.
- Shifley ET, Vanhorn KM, Perez-Balaguer A, Franklin JD, Weinstein M, Cole SE. 2008. Oscillatory lunatic fringe activity is crucial for segmentation of the anterior but not posterior skeleton. *Development* 135:899-908.
- Shinmyo Y, Mito T, Matsushita T, Sarashina I, Miyawaki K, Ohuchi H, Noji S. 2005. caudal is required for gnathal and thoracic patterning and for posterior elongation in the intermediate-germband cricket *Gryllus bimaculatus*. *Mech Dev* 122:231-239.
- Shum AS, Poon LL, Tang WW, Koide T, Chan BW, Leung YC, Shiroishi T, Copp AJ. 1999. Retinoic acid induces down-regulation of Wnt-3a, apoptosis and diversion of tail bud cells to a neural fate in the mouse embryo. *Mech Dev* 84:17-30.
- Sirbu IO, Duester G. 2006. Retinoic-acid signalling in node ectoderm and posterior neural plate directs left-right patterning of somitic mesoderm. *Nat Cell Biol* 8:271-277.
- Soriano P. 1997. The PDGF alpha receptor is required for neural crest cell development and for normal patterning of the somites. *Development* 124:2691-2700.
- Stauber M, Sachidanandan C, Morgenstern C, Ish-Horowicz D. 2009. Differential axial requirements for lunatic fringe and Hes7 transcription during mouse somitogenesis. *PLoS One* 4:e7996.
- Stavridis MP, Collins BJ, Storey KG. 2010. Retinoic acid orchestrates fibroblast growth factor signalling to drive embryonic stem cell differentiation. *Development* 137:881-890.
- Takemoto T, Uchikawa M, Yoshida M, Bell DM, Lovell-Badge R, Papaioannou VE, Kondoh H. 2011. Tbx6-dependent Sox2 regulation determines neural or mesodermal fate in axial stem cells. *Nature* 470:394-398.
- Tam PP, Beddington RS. 1987. The formation of mesodermal tissues in the mouse embryo during gastrulation and early organogenesis. *Development* 99:109-126.
- Tenin G, Wright D, Ferjentsik Z, Bone R, McGrew MJ, Maroto M. 2010. The chick somitogenesis oscillator is arrested before all paraxial mesoderm is segmented into somites. *BMC Dev Biol* 10:24.
- Tortelote GG, Hernandez-Hernandez JM, Quaresma AJ, Nickerson JA, Imbalzano AN, Rivera-Perez JA. 2013. Wnt3 function in the epiblast is required for the maintenance but not the initiation of gastrulation in mice. *Dev Biol* 374:164-173.
- Tschopp P, Tarchini B, Spitz F, Zakany J, Duboule D. 2009. Uncoupling time and space in the collinear regulation of Hox genes. *PLoS Genet* 5:e1000398.
- Tzouanacou E, Wegener A, Wymeersch FJ, Wilson V, Nicolas JF. 2009. Redefining the progression of lineage segregations during mammalian embryogenesis by clonal analysis. *Dev Cell* 17:365-376.
- van de Ven C, Bialecka M, Neijts R, Young T, Rowland JE, Stringer EJ, Van Rooijen C, Meijlink F, Novoa A, Freund JN, Mallo M, Beck F, Deschamps J. 2011. Concerted involvement of Cdx/Hox genes and Wnt signaling in morphogenesis of the caudal neural tube and cloacal derivatives from the posterior growth zone. *Development* 138:3451-3462.
- 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, Dolle P. 2005. Retinoic acid controls the bilateral symmetry of somite formation in the mouse embryo. *Science* 308:563-566.
- Vidigal JA, Morkel M, Wittler L, Brouwer-Lehmitz A, Grote P, Macura K, Herrmann BG. 2010. An inducible RNA interference system for the functional dissection of mouse embryogenesis. *Nucleic Acids Res* 38:e122.

- Vlangos CN, Siuniak AN, Robinson D, Chinnaiyan AM, Lyons RH, Jr., Cavalcoli JD, Keegan CE. 2013. Next-generation sequencing identifies the Danforth's short tail mouse mutation as a retrotransposon insertion affecting *Ptfl1a* expression. *PLoS Genet* 9:e1003205.
- Wahl MB, Deng C, Lewandoski M, Pourquie O. 2007. FGF signaling acts upstream of the NOTCH and WNT signaling pathways to control segmentation clock oscillations in mouse somitogenesis. *Development* 134:4033-4041.
- Wilson V, Beddington R. 1997. Expression of T protein in the primitive streak is necessary and sufficient for posterior mesoderm movement and somite differentiation. *Dev Biol* 192:45-58.
- Wilson V, Olivera-Martinez I, Storey KG. 2009. Stem cells, signals and vertebrate body axis extension. *Development* 136:1591-1604.
- Yamaguchi TP, Harpal K, Henkemeyer M, Rossant J. 1994. *fgfr-1* is required for embryonic growth and mesodermal patterning during mouse gastrulation. *Genes Dev* 8:3032-3044.
- Yamaguchi TP, Takada S, Yoshikawa Y, Wu N, McMahon AP. 1999. T (Brachyury) is a direct target of *Wnt3a* during paraxial mesoderm specification. *Genes Dev* 13:3185-3190.
- Yoshikawa Y, Fujimori T, McMahon AP, 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 JE, van de Ven C, Bialecka M, Novoa A, Carapuco M, van Nes J, de Graaff W, Duluc I, Freund JN, Beck F, Mallo M, Deschamps J. 2009. *Cdx* and *Hox* genes differentially regulate posterior axial growth in mammalian embryos. *Dev Cell* 17:516-526.
- Zhao X, Duyster G. 2009. Effect of retinoic acid signaling on *Wnt/beta-catenin* and *FGF* signaling during body axis extension. *Gene Expr Patterns* 9:430-435.

1A

Chapter 1

Introduction part B

At the start of colinear Hox gene expression: *cis*-features and *trans*-factors orchestrating the initial phase of Hox cluster activation

Roel Neijts¹ and Jacqueline Deschamps¹

¹Hubrecht Institute, Developmental Biology and Stem Cell Research, Uppsalalaan 8, 3584 CT Utrecht, and UMC Utrecht

Adapted from:

Neijts R and Deschamps J. At the start of colinear Hox gene expression: *cis*-features and *trans*-factors orchestrating the initial phase of Hox cluster activation,

Review to appear in the special issue of Developmental Biology (Elsevier) on the centennial of the Hubrecht Institute, (scheduled for 2017)

Abstract

Hox genes are well-known players in the generation and patterning of the vertebrate trunk and posterior body during embryogenesis. Their initial expression takes place shortly after the establishment of the primitive streak, in the posterior-most part of the mouse embryo and is a determinant step for setting up the definitive Hox expression boundaries along the antero-posterior body axis. The developmental signals and epigenetic mechanisms underlying this early activation remained unsolved until recently. The development of novel embryo-derived model systems, combined with methods that examine chromatin status and chromosome conformation, led to advanced understanding of the process of Hox activation in the early embryo. We here summarize how the early *cis*-regulatory Hox landscape becomes active upon receiving the right developmental signal, and we discuss the importance of the local topological segmentation of the HoxA cluster during early Hox activation.

Introduction

One of the most fascinating gene regulatory processes in developmental biology is the onset of temporal and spatial colinear expression of Hox genes (Krumlauf, 1994; Kmita and Duboule, 2003). Each of the four mammalian Hox clusters – A, B, C and D – starts to be activated at its 3' side. The process gradually extends to the middle of the clusters in register with developmental time, until it reaches the 5'-most genes, the Hox13 paralogs. The colinear relationship between the position of a given Hox gene on the chromosome and its spatial domain of activation was first observed in *Drosophila* (Lewis, 1978). Since then, spatially colinear expression of Hox genes has been demonstrated to be widespread in the animal kingdom, whereas temporal colinearity was shown to be restricted to bilaterians that have maintained their Hox clusters in a relatively intact organization (Duboule, 2007; Noordermeer and Duboule, 2013). The sequential turning on of Hox genes over developmental time provides precursors of embryonic tissues with a position-specific 'Hox code' (Kessel and Gruss, 1991) along the trunk axis and along appendicular axes such as the limb or external genitals. Timing of initial Hox gene expression is intimately linked to the later spatial expression domains. These domains – and therefore the timing – of Hox gene expression are crucial for normal embryonic development, and ectopically expressed Hox genes cause severe developmental abnormalities. Striking examples of failure to obey this requirement are the suppressed rib formation by ectopic *Hoxa10* (Vinagre et al., 2010), the formation of lumbar ribs by ectopic *Hoxb6* (Vinagre et al., 2010), and premature arrest of posterior axial growth by precociously expressed Hox13 genes (Young et al., 2009).

Here we focus on the initial onset of Hox gene expression that shortly follows the specification of the primary body axis. The emergence of novel model systems to study early embryogenesis, like epiblast stem cells (EpiSCs) (Brons et al., 2007; Tesar et al., 2007), gastruloids (van den Brink et al., 2014) and other embryonic stem cell (ESC)-derived systems (Henrique et al., 2015; Etoc et al., 2016), accelerated our understanding of the cellular, genetic and epigenetic aspects of the regulation of early developmental genes like Hox genes. These new models combined with the application of chromosome conformation capture-based technologies, and methods that examine the chromatin status (like ChIP-seq and ATAC-seq) revealed the molecular genetic mechanisms that were thus far not easy to explore in early embryos. We discuss the developmental signals and epigenetic events that are at the basis of the transcriptional initiation of the earliest Hox genes to be expressed. We describe the interplay between *trans*-acting factors and *cis*-regulatory elements of the segmented Hox landscape guiding progression of Hox gene expression towards the precursors of the tissues that will generate the vertebrate axial structures.

1. Early inducing signals and initial Hox activation in the gastrulating embryo

Priming, initiation and spreading of early Hox gene transcription

In the mouse embryo, gastrulation starts in the proximo-posterior epiblast, at a site demarcated by expression of *T Brachyury* and *Wnt3* at embryonic day (E)6.2 (Rivera-Perez and Magnuson, 2005). During early gastrulation the primitive streak gradually extends towards the distal tip of the embryo that has the shape of an egg cylinder. The cells that first ingress through the streak do not contribute to the embryo proper, but to extraembryonic tissues among which the allantois (Lawson et al., 1991). The first Hox-positive region in the embryo is the very posterior part of the fully extended primitive streak at E7.2 (Deschamps and Wijgerde, 1993; Gaunt and Strachan, 1994; Forlani et al., 2003). It was shown that the earliest Hox gene is primed for expression, one full day before transcripts can be detected by in situ hybridization (Forlani et al., 2003). After the 3'-most Hox gene is turned on in the posterior streak area, its expression domain spreads anterior-wards by a process that does not involve cell migration (Deschamps and Wijgerde, 1993; Forlani et al., 2003) to reach the anterior part of the streak that since recently is known to harbor axial progenitors among which long term bipotent neuromesodermal progenitors (NMPs) (Wilson et al., 2009; Wymeersch et al., 2016). More 5'-located Hox genes start to be transcribed in the posterior streak area subsequently to the 3' genes. Their expression domain in its turn spreads anteriorly. The anterior streak region, corresponding to the NMP-containing growth zone from which the trunk axial tissues will be formed, thus sequentially expresses a more and more posterior Hox code (Figure 1) (Deschamps and van Nes, 2005). Although this first wave of Hox expression does not yet concern differentiated tissues from the definitive embryonic germ layers, timing defects at this early stage of Hox transcription results in phenotypic abnormalities. Deletion of an early *Hoxc8* enhancer, which is active during the initial phase of the gene expression, results in homeotic transformations along the vertebral column later on (Juan and Ruddle, 2003). The Hox expression domains later extend more anteriorly than the node and reach their spatial boundaries in the paraxial mesoderm and, independently, in the neurectoderm (Deschamps and van Nes, 2005) (Forlani et al., 2003). The precise timing of early Hox initiation in the primitive streak is a first and determinant step for the later setting up of the Hox expression boundaries in the embryo.

Hox initiation signals: Wnt ligands and other proposed candidates

Several developmental signals have been shown to influence early Hox gene expression in the embryo. They include canonical Wnt ligands, Fgf and Gdf11 (Liu et al., 2001; Deschamps and van Nes, 2005; Gaunt et al., 2013). Retinoic acid (RA) is regulating Hox gene expression in the hindbrain (Gavalas and Krumlauf, 2000) but its role in the earliest transcriptional initiation of Hox genes in the early embryo is unlikely. The next sections will discuss the modalities of action of these signals on Hox gene expression.

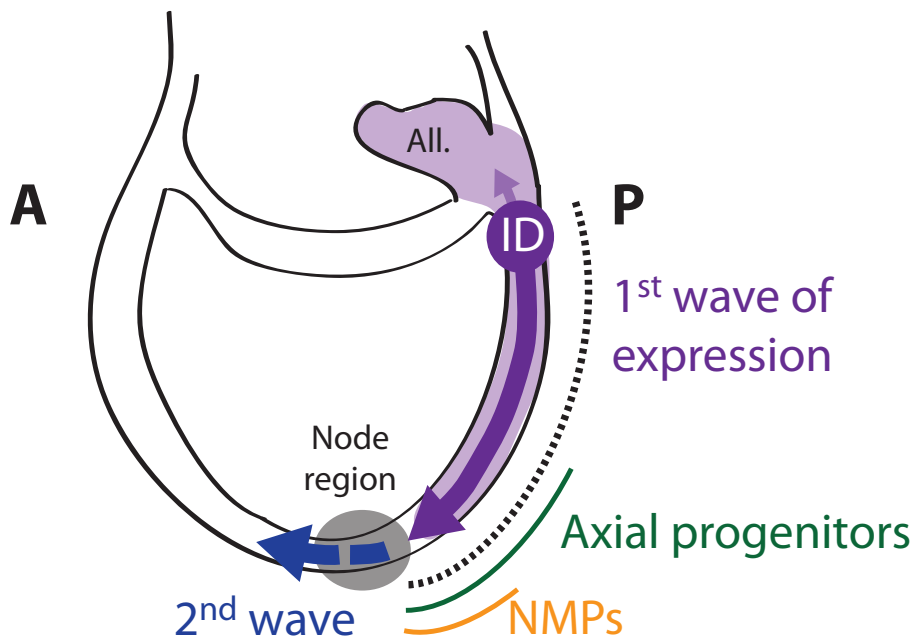


Figure 1: Schematic representation of the two waves of Hox expression in the early posterior embryo. *Hoxa1* expression (in purple) starts in the posterior-most part of the streak (ID, induction domain) at E7.2. From the ID, cells expressing Hox contribute to extraembryonic tissue (the allantois, all.; small purple arrow) and *Hoxa1* expression spreads anterior-wards towards (large purple arrow). This spreading signal reached the anterior streak and node region, which harbor axial progenitors including NMPs. As a consequence these progenitors are sequentially expressing a more and more posterior Hox code. In the second wave (blue arrow), the Hox code obtained in the node is regulated independently in neuroectodermal and mesodermal tissue; this code definitely patterns the tissue along the antero-posterior axis. The primitive streak is indicated by a dashed line. A, anterior. P, posterior.

Wnt as instructive Hox initiation signal

A role for Wnt/ β -catenin in antero-posterior polarity of the body axis predates the bilaterian ancestor, since a Wnt-dependence in the patterning of the primary axis is observed amongst cnidarians (Petersen and Reddien, 2009). In mouse embryos, early anterior-posterior regionalization is manifested by asymmetrical expression of *Wnt3* (Rivera-Perez and Magnuson, 2005), which is restricted to the proximal and posterior epiblast. *Wnt3* has been shown to be essential for the formation of the primitive streak and gastrulation. *Wnt3* mutant embryos do not express any Hox genes (Liu et al., 1999). Involvement of Wnt signals in the initiation of 3' Hox genes was suggested by Forlani et al., 2003. Pre-gastrulation embryos exposed to the Wnt agonist Chiron precociously express *Hoxa1* and *Hoxb1* (see Chapter 4 of this thesis) (Neijts et al., 2016) demonstrating the early responsiveness of 3' Hox genes to Wnt signals. Epiblast stem cells (EpiSCs) generated from wild type or *Wnt3* mutant epiblast start to express Hox genes as soon as Wnt signaling is present. Moreover, deletion of Wnt-sensitive enhancers in

the 3' regulatory landscape of *HoxA* could prevent Wnt-induced activation of *Hoxa1* (Chapter 4).

The signal inherent to posteriorization in bilaterians – the Wnt-dependent stabilization of β -catenin – thus functions as a master regulator to initiate Hox gene expression at the right time in the embryo. It allows the Hox transcription domains to reach the axial progenitor region at the time, when the hindbrain to caudal-most tissues are laid down in the posterior embryonic growth zone. Interestingly, it has been hypothesized that the Hox gene regulation was co-opted by the Wnt regulatory network prior to the last common ancestor of cnidarian and bilaterian animals (Ryan and Baxevanis, 2007). From that co-option on, Wnt signaling and Hox expression would be at work in concert in a posterior genetic network underlying body extension and patterning.

Is RA required for Hox initiation?

Numerous lines of evidence, *in vivo* and in embryo-derived model systems, have shown that retinoids are able to influence Hox gene expression during embryogenesis (Gavalas and Krumlauf, 2000; Oosterveen et al., 2003). RA is a well-known inducer of colinear Hox gene expression in pluripotent mouse embryonic carcinoma cells, embryonic stem cells (ESCs) and human pluripotent cells (Colberg-Poley et al., 1985; Breier et al., 1986; Chambeyron and Bickmore, 2004; Agger et al., 2007). Despite the ability of RA to induce Hox genes in pluripotent cell lines and its effect on the Hox code in late developing embryos, the role of endogenous retinoids in the initial activation of the clusters in vertebrates remains elusive.

In murine early somite embryos, significant levels of retinoids have been found in the node region (Hogan et al., 1992). RA levels were not found where and when Hox gene transcription is primed and initiated: in the posterior part of the primitive streak. Embryos deficient for the RA-producing enzyme *Aldh1a2* (or: *Raldh2*) could still initiate the 3' Hox gene *Hoxb1* (Roelen et al., 2002) and were able to gastrulate and to generate some anterior somites (Niederreither et al., 1999). In addition, the deletion of a proximal RA-responsive element (RARE) at the *Hoxa1* locus affects the gene expression level at E7.5, but it does not prevent the initial transcriptional induction (Dupe et al., 1997). The deletion of a RARE 3' near *Hoxb1* leads to the absence of *Hoxb1* transcription in the posterior part of the E8.25 embryos, but the situation was not examined at the time of endogenous *Hoxb1* initiation (around E7.2) (Marshall et al., 1994).

In summary, the Hox clusters are responsive to exogenous RA signals from early stages on. But despite this RA sensitivity and the fact that retinoids induce colinear Hox gene expression in pluripotent cell lines, a role for endogenous RA in initial Hox induction remains unestablished.

Are Cdx genes involved in Hox transcriptional initiation?

Cdx and Hox genes are evolutionarily linked as they both derive from an ancient protoHox gene or gene cluster (Pollard and Holland, 2000). The similar *in vivo* expression dynamics of

3' Hox and Cdx genes are obvious: both gene subfamilies start to be expressed in the posterior primitive streak around E7.2 and their transcription similarly spreads anterior-wards (Young and Deschamps, 2009). Like the 3' Hox genes, *Cdx2* is initially induced by Wnt signals in epiblast-derived EpiSCs (Chapter 3)(Amin et al., in press). Cdx genes were found to regulate trunk Hox genes and to accordingly modify the identity of axial trunk tissues (Shashikant and Ruddle, 1996; Charite et al., 1998; Gaunt et al., 2004; Tabaries et al., 2005). Cdx mutants that over-express trunk Hox genes were rescued in their posterior truncation phenotype (Young et al., 2009); both Cdx and Hox genes are involved in the maintenance of NMPs during axial elongation (see Chapter 1, part A and Chapter 3).

Our laboratory recently identified the direct downstream targets of *Cdx2*, including Hox genes, by ChIP-seq experiments in Wnt-stimulated EpiSCs (Chapter 3). We have found that Cdx is required for making the DNA accessible at several *cis*-elements within the middle part the Hox clusters (Chapter 5)(Neijts et al., unpublished). In contrast, the chromatin at the 3' parts of the HoxA and HoxB clusters is open and decorated by active histone mark H3K27ac independently of Cdx. These findings point to a function of Cdx during the activation of the middle/trunk Hox genes in the post-initiated clusters (Chapter 5). In accordance, Cdx triple mutant embryos could still initiate the early *Hoxb1* gene like wild type embryos do (Chapter 2) (van Rooijen et al., 2012). Therefore it is unlikely that Cdx transcription factors are involved in the initial transcription of 3' Hox genes. Rather they control the colinear activation of the subsequent clustered genes (see Chapter 2 and 5).

2. Dissecting the Hox locus: *cis*-regulatory modules

The Hox regulatory landscape is subdivided over topological domains

The expression of a developmental gene is regulated by its *cis*-regulatory landscape consisting of enhancers, insulators and other architectural elements which can be located proximal to the gene or dispersed over large genomic distances [reviewed by (Spitz, 2016)]. Besides numerous regulatory elements within the Hox clusters themselves, the very large gene-poor desert regions that flank the clusters on both sides – in particular HoxA and HoxD – harbor various elements that influence Hox spatiotemporal expression during development (Montavon and Duboule, 2013). Over the years, the long-range regulatory potential of these genomic landscapes has been studied intensively using a comprehensive collection of mouse strains carrying genomic rearrangements (Tschopp and Duboule, 2014). More recently genome-wide approaches such as ChIP-seq and ATAC-seq (Buenrostro et al., 2013) made it possible to identify and map different *cis*-regulatory elements in and around the clusters.

In addition to proximal and distal *cis*-regulatory sequences and their chromatin, the three-dimensional organization and the spatial compartmentalization of loci play a major role in developmental gene regulation (de Laat and Duboule, 2013). Initially, DNA-FISH has

been the most pertinent method to study the physical architecture of the Hox loci. Using this method it was observed that during HoxB activation, transcriptionally induced Hox genes could loop out of their ‘chromosome territory’ (Chambeyron and Bickmore, 2004). This chromatin dynamics of HoxB was first observed in RA-induced ESCs, and later confirmed in gastrulating embryos (Chambeyron et al., 2005). The introduction of chromosome conformation capture-based approaches, like 3C, 4C-seq and HiC [reviewed by (Denker and de Laat, 2016)] in developmental biology allowed the in-depth study of the genome architecture, chromatin compartments and gene-enhancer contacts – and their dynamics – over developmental time. Using 4C-seq, the Duboule laboratory observed that the activated genes shift from an inactive chromatin domain to an active domain during colinear activation of Hox gene expression between E8.5 and E12.5 in tissues along the embryonic axis, (Noordermeer et al., 2011). DNA-FISH and 4C-seq independently established the concept that the conformation of Hox loci is very dynamic during progressive Hox gene activation, and that higher order re-arrangements take place over time as colinear Hox expression continues.

The discovery of the partition of chromosomes into large segments called ‘topologically associating domains’ (TADs) (Dixon et al., 2012; Nora et al., 2012; Sexton et al., 2012) was very important to understand the regulatory logic of the Hox *cis*-landscape. The HoxA and HoxD clusters were found to lie at the junction of two topological domains (Dixon et al., 2012). It is this boundary position that probably allowed a bimodal and stepwise regulation to control the expression of 3’ Hox genes separately from their 5’ neighbors during mouse limb outgrowth and possibly during axial development as well (Andrey et al., 2013; Darbellay and Duboule, 2016). Such a TAD-based bipartite regulatory mode of subsets of genes was not seen for the HoxB and HoxC clusters (Dixon et al., 2012). Interestingly, these latter clusters have an inherent *cis*-organization from which the 3’ and 5’ Hox genes are already physically separated: in HoxB, *Hoxb13* is located so remotely that it is isolated *in cis* from the *Hoxb1-Hoxb9* genes, and HoxC lacks the 3’ part of the cluster.

The distribution of the HoxA and HoxD clusters over two different TADs originated in an ancestor in the vertebrate lineage. In snakes and teleosts the bipartite architecture is present (Woltering et al., 2014; Guerreiro et al., 2016), whereas only a single (3’) TAD can be discerned on the Hox locus of the early chordate amphioxus (Acemel et al., 2016). It might well be that the origin of the 5’ TAD – with its regulatory influence predominantly over the 5’-most Hox genes – is linked to the appearance of novel body appendages at the root of the vertebrate lineage, like the fin. Functional regulatory elements have been found in the 5’ flanking Hox domain of teleosts (zebrafish) and more primitive bony fish (the spotted gar) (Gehrke and Shubin, 2016). Interestingly, the evolutionarily conserved 5’ elements from the spotted gar, active in the fin, drive gene expression in the distal limb (Gehrke et al., 2015) similarly to the mouse elements residing in the HoxA and HoxD 5’ TADs (Montavon et al., 2011; Andrey et al., 2013; Berlivet et al., 2013). A functional second Hox TAD therefore correlates with evolutionary invention of a secondary body axis. It is not ruled out that the bimodal regulation alternatively might have arisen for controlling the vertebrate primary axis, since it

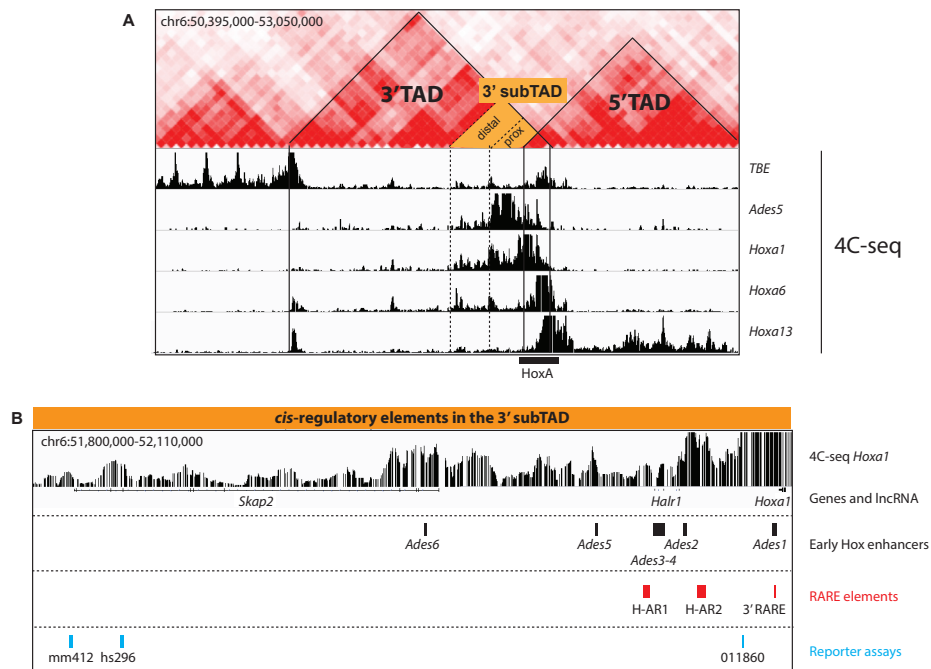


Figure 2: The *cis*-architecture of the *HoxA* locus includes a 3' regulatory subTAD. A) The global *cis*-conformation of the *HoxA* locus, which is located at the boundary domain of a 3' TAD and a 5' TAD (HiC data from Dixon et al., 2012). 4C-seq viewpoints are taken from the *TBE* element at the 3' boundary of the 3' TAD, from the *Ades5* enhancer and *Hoxa1* within the 3' subTAD (which is indicated in yellow), and from *Hoxa6* and *Hoxa13*. Depending on its position in the locus, favored interactions take place in different TADs. B) The 3' flanking region of the *HoxA* cluster (in the 3' subTAD) comprises 3' Hox genes (*Hoxa1* is shown), lncRNA *Halr1* and *Skap2*. Regulatory functions of *Ades* enhancers are described in Chapter 4. *Halr1* functions and dynamics are described in Guttman et al., 2011; Maamar et al., 2013; De Kumar et al., 2015; Yin et al., 2015; Liu et al., 2016. The early 3' RARE was described first in Langston and Gudas, 1992. H-AR1 and H-AR2 RARE-rich domains are described in De Kumar et al., 2015. Finally, public databases VISTA (Visel et al., 2007) and TRACER (Chen et al., 2013) revealed additional regulatory units (mm412 and hs296, VISTA) and regulatory sensitivity respectively (011860, TRACER). The BAC domains of Nolte et al., 2013 are not included. Moreover, the *TBE* region (1 Mb away from *Hoxa1*) gives a posterior trunk Hox-like activity domain as well (Chapter 5)(not shown in this figure).

is primordial to tightly and differentially regulate early (anterior) versus later posterior Hox genes. Whatever it may be, elements involved in late body axis regulation have not yet been identified in the 5' TAD.

The early Hox *cis*-regulatory landscape

Recently the molecular events underlying the earliest onset of Hox gene expression were studied, using Wnt-stimulated epiblast-derived EpiSCs representing the posterior post-implantation epiblast (Chapter 4)(Neijts et al., 2016). We could appreciate a bipartite distribution of contacts between genes from the *HoxA* cluster located within the intersection of the two large

TADs, by 4C-seq from different viewpoints (Figure 2A). The 3' HoxA genes interact mostly with the 3' TAD. HoxA genes at the center of the cluster mainly have contacts within the cluster itself and in the 3' TAD. The 5'-most *Hoxa13* gene interacts with the opposite (5') TAD, as previously seen at later stages in the limb (Berlivet et al., 2013; Lonfat et al., 2014) (Figure 2A).

More specifically, the 3' part of the cluster (*Hoxa1-Hoxa3*) highly interacts with a proximal region including the neighboring gene *Skap2* (Chapter 4). This region was identified as being a 3' subTAD (Figure 2). Before actual Hox gene transcription, during the induced transition from ESCs to EpiSCs, this 3' segment becomes more compact (Chapter 4). In this subTAD, several Wnt responsive ('Ades') enhancers were identified, which were demonstrated to be active in the posterior-most part of the streak at the moment of Hox initiation. The compact conformation of the HoxA early *cis*-landscape, forming a segment together with the 3' side of the cluster, and the presence of multiple Wnt responsive enhancers result in a *cis*-environment that is primed for gene expression. The chromatin organization of the 3' neighborhood of HoxA is such that the cluster exclusively can be activated on that side in response to incoming Wnt signals that are inherent to the posterior specification of the early embryo. The trunk Hox genes and the latest *Hoxa13* gene lack intense contacts with the early subTAD and are thus isolated from and insensitive to the Wnt-activated enhancers located the subTAD.

Besides the above described enhancers, the 3' flanking region of HoxA harbors additional regulatory information. Krumlauf and colleagues investigated a 40 kb-large portion of the proximal HoxA 3' subTAD (Nolte et al., 2013). The proximal-most 10 kb, containing the *Ades1* enhancer, contains a 3' RARE (Langston and Gudas, 1992; Dupe et al., 1997), and produces a pattern that is very similar to endogenous *Hoxa1* expression (Nolte et al., 2013). Moreover, the 3' subTAD harbors the long non-coding RNA (lncRNA) *Halr1* that resides in the *Heater* locus (De Kumar et al., 2015; De Kumar and Krumlauf, 2016). Several groups have been dissecting the function of this region in ESCs (Guttman et al., 2011; Maamar et al., 2013; De Kumar et al., 2015; Yin et al., 2015; Liu et al., 2016). *Halr1* acts as a transcriptional enhancer and is activated upon RA exposure. Interestingly, this lncRNA binds the HoxA locus *in trans* to repress gene expression by preventing H3K27me3 demethylation (Yin et al., 2015; Liu et al., 2016). Although *Halr1* was shown to be important for proper activation of the HoxA locus in ESCs, mice lacking the lncRNA had no reported developmental abnormalities (Sauvageau et al., 2013; Lai et al., 2015), suggesting a role for *Halr1* in later fine-tuning rather than in any crucial early transcriptional control. Besides different isoforms of *Halr1* from both strands, unspliced short transcripts are also produced from the *Heater* region (De Kumar et al., 2015). As *Ades3-4* and *Ades2* enhancers are located in this interval, they could be responsible for the transcription of 'enhancer-RNA' (eRNA) (Li et al., 2016) that include unspliced short fragments.

In Figure 2B we have summarized the most important regulatory elements which are identified in the early regulatory landscape flanking HoxA. The compact 3' subTAD is very rich in regulatory regions which are involved in both the initiation and in the later regulation

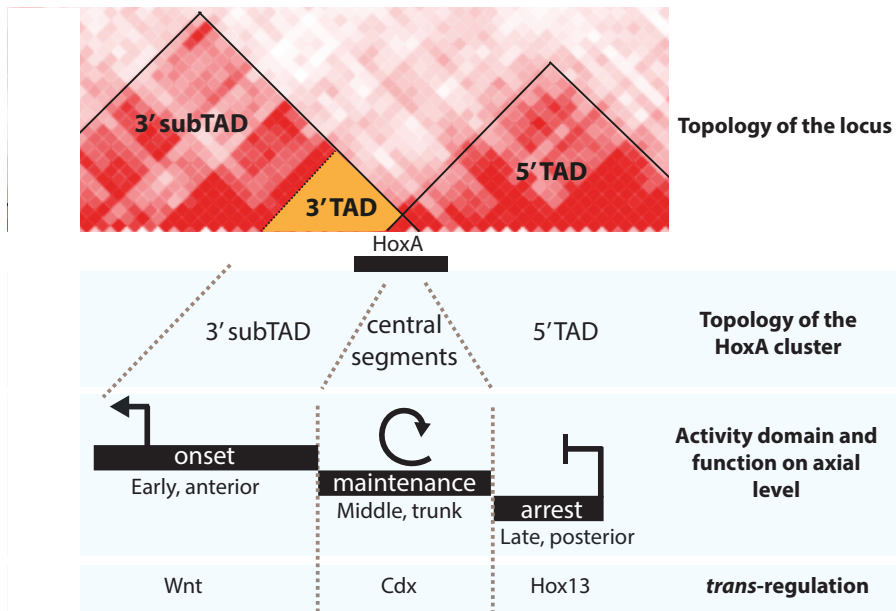


Figure 3: The *cis* and *trans* characteristics of the HoxA locus during axial elongation. Upper panel: the HoxA locus is subdivided over a 3' TAD, that included a proximal 3' subTAD (in yellow), and a 5' TAD (data from Dixon et al. 2012). Lower panels: The HoxA cluster is compartmentalized in three segments: the 3' subTAD, a central part, and a 5' TAD. The 3' subTAD is initiated first, by Wnt, and regulates the anterior-most axial tissues. The central HoxA part is controlled by Cdx transcription factors and is involved in the maintenance of axial progenitors and in the segmentation of the trunk. The 5' part of HoxA, harboring *Hoxa13*, is expressed relatively late and is responsible for axial arrest.

of the HoxA cluster. Whether Ades enhancer functions and lncRNA *Halr1* activity are overlapping in modulating developmental processes needs to be elucidated. In any case, the absence of *Halr1* transcripts in differentiated ESCs expressing the nascent mesodermal marker *T Brachyury* (Yin et al., 2015), suggests that *Halr1* and the Ades enhancers follow different regulatory rules and answer to different regulatory input.

The topological segmentation of Hox restricts regulatory input in time and space

Besides a 3' subTAD we could identify additional sub-segments that divide the HoxA region (see Figure 3). Adjacent to the 3' region, a middle segment was identified in which *cis*-elements are dependent on Cdx transcription factors for their accessibility (Chapter 5). Since expression of Cdx genes is induced by Wnt as well, this transcription factors inherently function as secondary activators. Lastly, a 5' segment exclusively contains *Hoxa13* and its associated enhancers (Chapter 4). Upon Wnt-stimulation of EpiSCs we observed that the cluster topology becomes reorganized around the boundaries of these segments (Chapter 4).

The compartmentalization of the HoxA cluster allows a segmental-wise – and relatively independent – activation of different sets of Hox genes and it has important regulatory implications. This could serve as a good strategy to restrict regulatory influence to only specif-

ic subsets of Hox genes. The evolutionarily emergence of Wnt- and RA-responsive enhancers and lncRNAs in the HoxA 3' subTAD – inside or outside of the cluster – primarily affects the 3' genes *in cis*. Without deregulation of middle or more posterior Hox genes, the 3' subTAD may have formed an evolutionary playground for early and anterior regulation. In the topologically isolated 3' subTAD a diversity of transcriptional enhancers acting on developmental early and anterior axial tissues like the hindbrain, cranial neural crest cells and branchial arches could arise (Maconochie et al., 1999; Parker et al., 2014; McEllin et al., 2016; Chapter 4). In addition, the segmentation allows 3' enhancers outside of the 3' subTAD, like the very distant TBE (1 Mb away from *Hoxa1*) to exert their regulatory function independently of the 3' subTAD on HoxA cluster members other than the 3' genes: the trunk and posterior Hox genes (Chapter 5).

The molecular basis of the segmentation of the 3' TAD and the HoxA cluster is not understood. Binding of the structural protein CTCF is associated with the boundaries of TAD domains (Yaffe and Tanay, 2011; Dixon et al., 2012). Acting as a insulation protein, CTCF is involved in maintaining the Hox activity domains in ESC-derived motor neurons as deletion of intra-cluster CTCF sites results in an altered Hox code and undue spreading of chromatin decoration (Narendra et al., 2015). Whether CTCF plays a role in the stepwise activation of the Hox cluster in the early embryo is unknown. The initiation of Hox gene transcription is very likely independent of CTCF binding: in ESCs and in human pluripotent cell lines CTCF occupancy is mainly found at the 5' half of the cluster and not at the 3' end (Ferraiuolo et al., 2010; Narendra et al., 2015). The important CTCF site CBS5 between *Hoxa7* and *Hoxa9* (Kim et al., 2011) marks a topological boundary in ESCs (Dixon et al., 2012) and in EpiSCs, and demarcates the 5'-limit of the Cdx2 binding domain (Chapter 5). Therefore, CTCF might rather play a role during the post-initiation Hox regulation.

Final remarks

The topological organization of the HoxA cluster stands at the basis of a polarity in Hox cluster activation by incoming developmental signals. The battery of Wnt-responsive enhancers at the early side of the cluster elicit a robust transcriptional activity, the action of which is very locally restricted to the genes and elements present in the 3' subTAD. The 5' TAD does not contain early Wnt-responsive enhancers. Importantly, more distant genes in the middle segment are fully covered by repressive histone modifications and depend on Cdx-driven chromatin opening for activation.

The three Hox clusters with an intact 3' region (i.e. A, B and D) are responding to Wnt exposure by activating the transcription of their 3' genes in EpiSCs (see Chapter 4). However, they each probably use different or additional strategies to bring about 3'-polarized Wnt-dependent transcriptional initiation of the locus. For instance, the HoxB cluster presents a localized depletion of PcG-PRC1 Polycomb component Ring1b coverage at the 3'-most gene, *Hoxb1*. Upon Wnt-exposure, the *Hoxb1* region – and its 3' enhancers (Marshall et al., 1994; Nolte et al., 2013) – likely function as an exposed 'docking' region for transcriptional

machinery.

In conclusion, as illustrated earlier for other regulatory properties of the Hox clusters, evolution has modelled the 3' landscape and structural features of the Hox clusters in varying ways ending up with their sensitivity to transcriptional activating Wnt signals in the gastrulating embryo. This initial activation unleashes the cascade of events that will allow trunk and posterior tissues to be correctly laid down and patterned.

1B

References

- Acemel RD, Tena JJ, Irastorza-Azcarate I, Marletaz F, Gomez-Marin C, de la Calle-Mustienes E, Bertrand S, Diaz SG, Aldea D, Aury JM, Manganot S, Holland PW, Devos DP, Maeso I, Escriva H, Gomez-Skarmeta JL. 2016. A single three-dimensional chromatin compartment in amphioxus indicates a stepwise evolution of vertebrate Hox bimodal regulation. *Nat Genet* 48:336-341.
- Agger K, Cloos PA, Christensen J, Pasini D, Rose S, Rappsilber J, Issaeva I, Canaani E, Salcini AE, Helin K. 2007. UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* 449:731-734.
- Amin S, Neijts R, Simmini S, van Rooijen C, Tan S, Kester L, van Oudenaarden A, Creighton MP, Deschamps J. in press. Cdx and T Brachyury co-activate growth signaling in the embryonic axial progenitor niche *Cell Reports*.
- 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* 340:1234167.
- Berlivet S, Paquette D, Dumouchel A, Langlais D, Dostie J, Kmita M. 2013. Clustering of tissue-specific sub-TADs accompanies the regulation of HoxA genes in developing limbs. *PLoS Genet* 9:e1004018.
- Breier G, Bucan M, Francke U, Colberg-Poley AM, Gruss P. 1986. Sequential expression of murine homeo box genes during F9 EC cell differentiation. *EMBO J* 5:2209-2215.
- Brons IG, Smithers LE, Trotter MW, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett SK, Clarkson A, Ahrlund-Richter L, Pedersen RA, Vallier L. 2007. Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* 448:191-195.
- Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* 10:1213-1218.
- Chambeyron S, Bickmore WA. 2004. Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev* 18:1119-1130.
- Chambeyron S, Da Silva NR, Lawson KA, Bickmore WA. 2005. Nuclear re-organisation of the Hoxb complex during mouse embryonic development. *Development* 132:2215-2223.
- Charite J, de Graaff W, Consten D, Reijnen MJ, 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.
- Chen CK, Symmons O, Uslu VV, Tsujimura T, Ruf S, Smedley D, Spitz F. 2013. TRACER: a resource to study the regulatory architecture of the mouse genome. *BMC Genomics* 14:215.
- Colberg-Poley AM, Voss SD, Chowdhury K, Gruss P. 1985. Structural analysis of murine genes containing homoeo box sequences and their expression in embryonal carcinoma cells. *Nature* 314:713-718.
- Darbellay F, Duboule D. 2016. Topological Domains, Metagenes, and the Emergence of Pleiotropic Regulations at Hox Loci. *Curr Top Dev Biol* 116:299-314.
- De Kumar B, Krumlauf R. 2016. HOXs and lincRNAs: Two sides of the same coin. *Sci Adv* 2:e1501402.
- De Kumar B, Parrish ME, Slaughter BD, Unruh JR, Gogol M, Seidel C, Paulson A, Li H, Gaudenz K, Peak A, McDowell W, Fleharty B, Ahn Y, Lin C, Smith E, Shilatifard A, Krumlauf R. 2015. Analysis of dynamic changes in retinoid-induced transcription and epigenetic profiles of murine Hox clusters in ES cells. *Genome Res* 25:1229-1243.
- de Laat W, Duboule D. 2013. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* 502:499-506.
- Denker A, de Laat W. 2016. The second decade of 3C technologies: detailed insights into nuclear organization. *Genes Dev* 30:1357-1382.
- Deschamps J, van Nes J. 2005. Developmental regulation of the Hox genes during axial morphogenesis in the mouse. *Development* 132:2931-2942.
- Deschamps J, Wijgerde M. 1993. Two phases in the establishment of HOX expression domains. *Dev Biol* 156:473-

- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485:376-380.
- Duboule D. 2007. The rise and fall of Hox gene clusters. *Development* 134:2549-2560.
- Dupe V, Davenne M, Brocard J, Dolle P, Mark M, Dierich A, Chambon P, Rijli FM. 1997. In vivo functional analysis of the Hoxa-1 3' retinoic acid response element (3'RARE). *Development* 124:399-410.
- Etoc F, Metzger J, Ruzo A, Kirst C, Yoney A, Ozair MZ, Brivanlou AH, Siggia ED. 2016. A Balance between Secreted Inhibitors and Edge Sensing Controls Gastruloid Self-Organization. *Dev Cell* 39:302-315.
- Ferraiuolo MA, Rousseau M, Miyamoto C, Shenker S, Wang XQ, Nadler M, Blanchette M, Dostie J. 2010. The three-dimensional architecture of Hox cluster silencing. *Nucleic Acids Res* 38:7472-7484.
- Forlani S, Lawson KA, Deschamps J. 2003. Acquisition of Hox codes during gastrulation and axial elongation in the mouse embryo. *Development* 130:3807-3819.
- Gaunt SJ, Cockley A, Drage D. 2004. Additional enhancer copies, with intact cdx binding sites, anteriorize Hoxa-7/lacZ expression in mouse embryos: evidence in keeping with an instructional cdx gradient. *Int J Dev Biol* 48:613-622.
- Gaunt SJ, George M, Paul YL. 2013. Direct activation of a mouse Hoxd11 axial expression enhancer by Gdf11/Smad signalling. *Dev Biol* 383:52-60.
- Gaunt SJ, Strachan L. 1994. Forward spreading in the establishment of a vertebrate Hox expression boundary: the expression domain separates into anterior and posterior zones, and the spread occurs across implanted glass barriers. *Dev Dyn* 199:229-240.
- Gavalas A, Krumlauf R. 2000. Retinoid signalling and hindbrain patterning. *Curr Opin Genet Dev* 10:380-386.
- Gehrke AR, Schneider I, de la Calle-Mustienes E, Tena JJ, Gomez-Marin C, Chandran M, Nakamura T, Braasch I, Postlethwait JH, Gomez-Skarmeta JL, Shubin NH. 2015. Deep conservation of wrist and digit enhancers in fish. *Proc Natl Acad Sci U S A* 112:803-808.
- Gehrke AR, Shubin NH. 2016. Cis-regulatory programs in the development and evolution of vertebrate paired appendages. *Semin Cell Dev Biol* 57:31-39.
- Guerreiro I, Gitto S, Novoa A, Codourey J, Nguyen Huynh TH, Gonzalez F, Milinkovitch MC, Mallo M, Duboule D. 2016. Reorganisation of Hoxd regulatory landscapes during the evolution of a snake-like body plan. *Elife* 5.
- Guttman M, Donaghey J, Carey BW, Garber M, Grenier JK, Munson G, Young G, Lucas AB, Ach R, Bruhn L, Yang X, Amit I, Meissner A, Regev A, Rinn JL, Root DE, Lander ES. 2011. lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* 477:295-300.
- Henrique D, Abranches E, Verrier L, Storey KG. 2015. Neuromesodermal progenitors and the making of the spinal cord. *Development* 142:2864-2875.
- Hogan BL, Thaller C, Eichele G. 1992. Evidence that Hensen's node is a site of retinoic acid synthesis. *Nature* 359:237-241.
- Juan AH, Ruddle FH. 2003. Enhancer timing of Hox gene expression: deletion of the endogenous Hoxc8 early enhancer. *Development* 130:4823-4834.
- 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 YJ, Cecchini KR, Kim TH. 2011. Conserved, developmentally regulated mechanism couples chromosomal looping and heterochromatin barrier activity at the homeobox gene A locus. *Proc Natl Acad Sci U S A* 108:7391-7396.
- Kmita M, Duboule D. 2003. Organizing axes in time and space; 25 years of colinear tinkering. *Science* 301:331-333.
- Krumlauf R. 1994. Hox genes in vertebrate development. *Cell* 78:191-201.
- Lai KM, Gong G, Atanasio A, Rojas J, Quispe J, Posca J, White D, Huang M, Fedorova D, Grant C, Miloscio L, Droguett G, Poueymirou WT, Auerbach W, Yancopoulos GD, Frenthewey D, Rinn J, Valenzuela DM. 2015. Diverse Phenotypes and Specific Transcription Patterns in Twenty Mouse Lines with Ablated LincRNAs. *PLoS One* 10:e0125522.

- Langston AW, Gudas LJ. 1992. Identification of a retinoic acid responsive enhancer 3' of the murine homeobox gene Hox-1.6. *Mech Dev* 38:217-227.
- Lawson KA, Meneses JJ, Pedersen RA. 1991. Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development* 113:891-911.
- Lewis EB. 1978. A gene complex controlling segmentation in *Drosophila*. *Nature* 276:565-570.
- Li W, Notani D, Rosenfeld MG. 2016. Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nat Rev Genet* 17:207-223.
- Liu GY, Zhao GN, Chen XF, Hao DL, Zhao X, Lv X, Liu DP. 2016. The long noncoding RNA Gm15055 represses Hoxa gene expression by recruiting PRC2 to the gene cluster. *Nucleic Acids Res* 44:2613-2627.
- Liu JP, Laufer E, Jessell TM. 2001. Assigning the positional identity of spinal motor neurons: rostrocaudal patterning of Hox-c expression by FGFs, Gdf11, and retinoids. *Neuron* 32:997-1012.
- Liu P, Wakamiya M, Shea MJ, Albrecht U, Behringer RR, Bradley A. 1999. Requirement for Wnt3 in vertebrate axis formation. *Nat Genet* 22:361-365.
- Lonfat N, Montavon T, Darbellay F, Gitto S, Duboule D. 2014. Convergent evolution of complex regulatory landscapes and pleiotropy at Hox loci. *Science* 346:1004-1006.
- Maamar H, Cabili MN, Rinn J, Raj A. 2013. linc-HOXA1 is a noncoding RNA that represses Hoxa1 transcription in cis. *Genes Dev* 27:1260-1271.
- Maconochie M, Krishnamurthy R, Nonchev S, Meier P, Manzanares M, Mitchell PJ, Krumlauf R. 1999. Regulation of Hoxa2 in cranial neural crest cells involves members of the AP-2 family. *Development* 126:1483-1494.
- Marshall H, Studer M, Popperl H, Aparicio S, Kuroiwa A, Brenner S, Krumlauf R. 1994. A conserved retinoic acid response element required for early expression of the homeobox gene Hoxb-1. *Nature* 370:567-571.
- McEllin JA, Alexander TB, Tumpel S, Wiedemann LM, Krumlauf R. 2016. Analyses of fugu hoxa2 genes provide evidence for subfunctionalization of neural crest cell and rhombomere cis-regulatory modules during vertebrate evolution. *Dev Biol* 409:530-542.
- Montavon T, Duboule D. 2013. Chromatin organization and global regulation of Hox gene clusters. *Philos Trans R Soc Lond B Biol Sci* 368:20120367.
- Montavon T, Soshnikova N, Mascrez B, Joye E, Thevenet L, Splinter E, de Laat W, Spitz F, Duboule D. 2011. A regulatory archipelago controls Hox genes transcription in digits. *Cell* 147:1132-1145.
- Narendra V, Rocha PP, An D, Raviram R, Skok JA, Mazzoni EO, Reinberg D. 2015. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science* 347:1017-1021.
- Neijts R, Amin S, van Rooijen C, Deschamps J. unpublished. Cdx is a crucial player during colinear Hox activation and defines a trunk segment in the Hox cluster topology.
- Neijts R, Amin S, van Rooijen C, Tan S, Creyghton MP, de Laat W, Deschamps J. 2016. Polarized regulatory landscape and Wnt responsiveness underlie Hox activation in embryos. *Genes Dev* 30:1937-1942.
- Neijts R, Simmini S, Giuliani F, van Rooijen C, Deschamps J. 2014. Region-specific regulation of posterior axial elongation during vertebrate embryogenesis. *Dev Dyn* 243:88-98.
- Niederreither K, Subbarayan V, Dolle P, Chambon P. 1999. Embryonic retinoic acid synthesis is essential for early mouse post-implantation development. *Nat Genet* 21:444-448.
- Nolte C, Jinks T, Wang X, Martinez Pastor MT, Krumlauf R. 2013. Shadow enhancers flanking the HoxB cluster direct dynamic Hox expression in early heart and endoderm development. *Dev Biol* 383:158-173.
- Noordermeer D, Duboule D. 2013. Chromatin architectures and hox gene collinearity. *Curr Top Dev Biol* 104:113-148.
- Noordermeer D, Leleu M, Splinter E, Rougemont J, De Laat W, Duboule D. 2011. The dynamic architecture of Hox gene clusters. *Science* 334:222-225.
- Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J, Gribnau J, Barillot E, Bluthgen N, Dekker J, Heard E. 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 485:381-385.

- Oosterveen T, Niederreither K, Dolle P, Chambon P, Meijlink F, Deschamps J. 2003. Retinoids regulate the anterior expression boundaries of 5' Hoxb genes in posterior hindbrain. *EMBO J* 22:262-269.
- Parker HJ, Bronner ME, Krumlauf R. 2014. A Hox regulatory network of hindbrain segmentation is conserved to the base of vertebrates. *Nature* 514:490-493.
- Petersen CP, Reddien PW. 2009. Wnt signaling and the polarity of the primary body axis. *Cell* 139:1056-1068.
- Pollard SL, Holland PW. 2000. Evidence for 14 homeobox gene clusters in human genome ancestry. *Curr Biol* 10:1059-1062.
- Rivera-Perez JA, Magnuson T. 2005. Primitive streak formation in mice is preceded by localized activation of Brachyury and Wnt3. *Dev Biol* 288:363-371.
- Roelen BA, de Graaff W, Forlani S, Deschamps J. 2002. Hox cluster polarity in early transcriptional availability: a high order regulatory level of clustered Hox genes in the mouse. *Mech Dev* 119:81-90.
- Ryan JF, Baxeavanis AD. 2007. Hox, Wnt, and the evolution of the primary body axis: insights from the early-divergent phyla. *Biol Direct* 2:37.
- Sauvageau M, Goff LA, Lodato S, Bonev B, Groff AF, Gerhardinger C, Sanchez-Gomez DB, Hacisuleyman E, Li E, Spence M, Liapis SC, Mallard W, Morse M, Swerdel MR, D'Ecclesiss MF, Moore JC, Lai V, Gong G, Yancopoulos GD, Friendewey D, Kellis M, Hart RP, Valenzuela DM, Arlotta P, Rinn JL. 2013. Multiple knockout mouse models reveal lincRNAs are required for life and brain development. *Elife* 2:e01749.
- Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G. 2012. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* 148:458-472.
- Shashikant CS, Ruddle FH. 1996. Combinations of closely situated cis-acting elements determine tissue-specific patterns and anterior extent of early Hoxc8 expression. *Proc Natl Acad Sci U S A* 93:12364-12369.
- Spitz F. 2016. Gene regulation at a distance: From remote enhancers to 3D regulatory ensembles. *Semin Cell Dev Biol* 57:57-67.
- Tabaries S, Lapointe J, Besch T, Carter M, Woollard J, Tuggle CK, Jeannotte L. 2005. Cdx protein interaction with Hoxa5 regulatory sequences contributes to Hoxa5 regional expression along the axial skeleton. *Mol Cell Biol* 25:1389-1401.
- Tesar PJ, Chenoweth JG, Brook FA, Davies TJ, Evans EP, Mack DL, Gardner RL, McKay RD. 2007. New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* 448:196-199.
- Tschopp P, Duboule D. 2014. The genetics of murine Hox loci: TAMERE, STRING, and PANTHERE to engineer chromosome variants. *Methods Mol Biol* 1196:89-102.
- van den Brink SC, Baillie-Johnson P, Balayo T, Hadjantonakis AK, Nowotschin S, Turner DA, Martinez Arias A. 2014. Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells. *Development* 141:4231-4242.
- 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.
- Vinagre T, Moncaut N, Carapuco M, Novoa A, Bom J, Mallo M. 2010. Evidence for a myotomal Hox/Myf cascade governing nonautonomous control of rib specification within global vertebral domains. *Dev Cell* 18:655-661.
- Visel A, Minovitsky S, Dubchak I, Pennacchio LA. 2007. VISTA Enhancer Browser--a database of tissue-specific human enhancers. *Nucleic Acids Res* 35:D88-92.
- Wilson V, Olivera-Martinez I, Storey KG. 2009. Stem cells, signals and vertebrate body axis extension. *Development* 136:1591-1604.
- Woltering JM, Noordermeer D, Leleu M, Duboule D. 2014. Conservation and divergence of regulatory strategies at Hox Loci and the origin of tetrapod digits. *PLoS Biol* 12:e1001773.
- Wymeersch FJ, Huang Y, Blin G, Cambray N, Wilkie R, Wong FC, Wilson V. 2016. Position-dependent plasticity of distinct progenitor types in the primitive streak. *Elife* 5:e10042.
- Yaffe E, Tanay A. 2011. Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture. *Nat Genet* 43:1059-1065.

- Yin Y, Yan P, Lu J, Song G, Zhu Y, Li Z, Zhao Y, Shen B, Huang X, Zhu H, Orkin SH, Shen X. 2015. Opposing Roles for the lncRNA Haunt and Its Genomic Locus in Regulating HOXA Gene Activation during Embryonic Stem Cell Differentiation. *Cell Stem Cell* 16:504-516.
- Young T, Deschamps J. 2009. Hox, Cdx, and anteroposterior patterning in the mouse embryo. *Curr Top Dev Biol* 88:235-255.
- Young T, Rowland JE, van de Ven C, Bialecka M, Novoa A, Carapuco M, van Nes J, de Graaff W, Duluc I, Freund JN, Beck F, Mallo M, Deschamps J. 2009. Cdx and Hox genes differentially regulate posterior axial growth in mammalian embryos. *Dev Cell* 17:516-526.

1B

1B

Chapter 2

Evolutionarily conserved requirement of Cdx for post-occipital tissue emergence

Carina van Rooijen¹, Salvatore Simmini^{1*}, Monika Bialecka^{1*}, **Roel Neijts^{1*}**, Cesca van de Ven¹, Felix Beck² and Jacqueline Deschamps¹

**Equal contribution*

¹Hubrecht Institute, Developmental Biology and Stem Cell Research, Uppsalalaan 8, 3584 CT Utrecht, and UMC Utrecht

²Biochemistry Department, University of Leicester, Leicester, LE1 9HN, UK

van Rooijen C, Simmini S*, Bialecka M*, **Neijts R***, van de Ven C, Beck F and Deschamps J. (2012) Evolutionarily conserved requirement of Cdx for post-occipital tissue emergence, *Development* 139(14):2576-83

**Equal contribution*

Abstract

Mouse *Cdx* genes are involved in axial patterning and partial *Cdx* mutants exhibit posterior embryonic defects. We found that mouse embryos wherein all three *Cdx* genes are inactivated fail to generate any axial tissue beyond the cephalic and occipital primordia. Anterior axial tissues are laid down and well patterned in *Cdx* null embryos, and a 3' most *Hox* gene is initially transcribed and expressed in the hindbrain normally. Axial elongation abruptly stops at the post-occipital level in the absence of *Cdx*, as the posterior growth zone loses its progenitor activity. Exogenous *Fgf8* rescues the posterior truncation of *Cdx* mutants, and the spectrum of defects of *Cdx* null embryos matches that resulting from the loss of posterior *FgfR1* signaling. Our data argue for a main function of *Cdx* in enforcing trunk emergence beyond the *Cdx* independent cephalo-occipital region, and for a downstream role of *FgfR1* signaling in this function. *Cdx* requirement for the post-head section of the axis is ancestral since it takes place in arthropods as well.

Introduction

During gastrulation of the mouse embryo, progenitors for trunk and tail tissues are found in an orderly way in the epiblast flanking the primitive streak, from its more posterior extension to its rostralmost limit abutting the node (Tam and Beddington, 1987; Lawson et al., 1991; Kinder et al., 1999). While the progenitors along the posterior and middle streak levels are transiently delivering descendants to extraembryonic and embryonic mesoderm and do not leave residing cells after they have ingressed in the streak, the anteriormost level of the primitive streak harbors a self-renewing, stem cell-like population of progenitors that go on contributing cells to the elongating axial tissues until the end of axial growth by tissue addition (Cambray and Wilson, 2002; Cambray and Wilson, 2007; Wilson et al., 2009). Clonal analysis during embryogenesis provided the evidence that bipotent long-term neuro/mesodermal (LT N/M) progenitors contribute descendants to extended axial domains (Tzouanacou et al., 2009). These progenitors are likely to correspond to the stem cell-like axial progenitors shown by Cambray and Wilson to be present in the node-streak border and along the antero-lateral primitive streak at embryonic day (E) 8.5, and in the chordo-neural hinge later on (Cambray and Wilson, 2002; Cambray and Wilson, 2007).

The mouse has three Cdx transcription factor encoding genes, *Cdx1*, *Cdx2* and *Cdx4*, homolog to *Drosophila caudal*. Cdx loss of function was first discovered to impair axial elongation when *Cdx2* was inactivated (Chawengsaksophak et al., 1997). Null mutants for *Cdx1* and *Cdx4* are not compromised in their axis extension but they fail to complete their axial development when missing an active allele of *Cdx2* (van den Akker et al., 2002; van Nes et al., 2006; Savory et al., 2009; Young et al., 2009). *Cdx4* null mutants heterozygote for *Cdx2* (that we call *Cdx2/4* mutants) also suffer from limited allantoic vessel invasion in the chorionic ectoderm, and the allantois of *Cdx2* null mutants fails to grow out, preventing placental labyrinth ontogenesis and survival of the embryo beyond E10.5 (Chawengsaksophak et al., 2004; van Nes et al., 2006; van de Ven et al., 2011). Compound mutants for the different Cdx genes revealed redundancy between them in allowing embryonic tissues from the three germ layers to expand as development proceeds (van den Akker et al., 2002; Savory et al., 2009; Young et al., 2009; Savory et al., 2011; van de Ven et al., 2011). Histological and gene expression analyses combined with the fate mapping information on the progenitors of axial tissues in the mouse embryo (Cambray and Wilson, 2002; Cambray and Wilson, 2007; Tzouanacou et al., 2009) led to the conclusion that the Cdx mutations in *Cdx2/4* compound mutants affect tissue generation from progenitors residing along the primitive streak and its continuation in the tail bud without causing apoptosis (Young et al., 2009; van de Ven et al., 2011). Genetic analysis revealed that the axial extension defects of these mutants could be rescued by either a gain of function of Hox genes belonging to the middle part of the Hox clusters, or by expressing an activated form of the Wnt signaling effector Lef1 in the spatio-temporal window of Cdx expression (Young et al., 2009). The latter information and subsequent grafting experiments of the region harboring stem cell-like axial progenitors for trunk and tail tissues from *Cdx2/4* mutants into wild type recipients revealed that the Cdx mutations disable the surrounding

niche of these progenitors rather than the progenitors themselves (Bialecka et al., 2010). So far the impact of ablating all three Cdx genes had not been tested. Cdx genes start to be transcribed at E7.2 in the posterior primitive streak. In order to study embryogenesis in the total absence of Cdx activity from early on in the epiblast, we generated mouse embryos totally deprived of Cdx expression using *Cdx1* and *Cdx4* null alleles (Subramanian et al., 1995; van Nes et al., 2006), and the *Cdx2* conditional allele that we produced (Stringer et al., 2012), in combination with the epiblast-specific *Sox2Cre*. We show here, using mouse embryos wherein the three Cdx genes are inactivated, that the realm of action of Cdx encompasses and is restricted to the entire trunk and tail sections of the axis. Ablation of all three Cdx genes causes agenesis of the axial domain posterior to the occipital region, involving the three germ layers. The key role of Fgf signaling in axial elongation was demonstrated by the fact that Fgf restores tissue emergence and gene expression in the embryonic growth zone of Cdx mutants in whole embryo cultures.

Results

Absence of Cdx prevents the generation of trunk and tail tissues during embryogenesis

Cdx triple null mutants were generated with mice carrying null alleles for *Cdx1*, *Cdx2* and *Cdx4*, and a conditional allele of *Cdx2* (Young, 2009; Stringer et al., 2012), in combination with a *Sox2Cre* transgene (Hayashi et al., 2002) allowing *Cdx2* inactivation in the ICM-derived embryonic tissues. Following this strategy, the epiblast of the triple mutant is totally deprived of Cdx activity from the earliest stage on, and the embryo proper is absolutely Cdx null. Triple Cdx mutant embryos were recovered at the expected Mendelian frequency, but their generation required intensive efforts because of low breeding performance of *Cdx1*^{-/-} *Cdx2*^{+/-} *Sox2Cre* males that were used in the final cross (see Materials and Methods). We analysed 52 Cdx triple null embryos, 21 embryos with a genotype *Cdx1*^{-/-} *Cdx2*^{-/-} *Cdx4*^{+/-} that are indistinguishable from the Cdx triple null mutants, and a larger number of *Cdx2* null mutants and wild type controls.

Cdx triple mutant (referred to as Cdx null from here on) embryos at E7.5 are indistinguishable from wild types (Fig. S1A,B), except for the fact that their allantois fails to grow as it is the case in *Cdx2* null embryos (Chawengsaksophak et al., 2004; van de Ven et al., 2011). Therefore they did not survive beyond E10.5 because they never establish a placental labyrinth. They can be recovered at E10.5 but seem to have arrested their development earlier, as judged from the fact that they are growth retarded at this stage in comparison with controls (Fig. S1E,F). Cdx null mutants were severely posteriorly truncated, and the axial level of the last tissues generated was anterior to the forelimb buds (Fig. S1C-F). In addition, they exhibited an open neural tube, a condition that was manifest as well in the absence of *Cdx1* and *Cdx2* (Savory et al., 2011). The Cdx null embryos generate a maximum of 5 somites. At E8.5, 25 out of 25 mutants had 5 somites instead of the 8-10 generated by control E8.5 littermates. At E9.5, 7 out of 8 mutants had 5 somites whereas age-matched controls had 22 to 25 somites

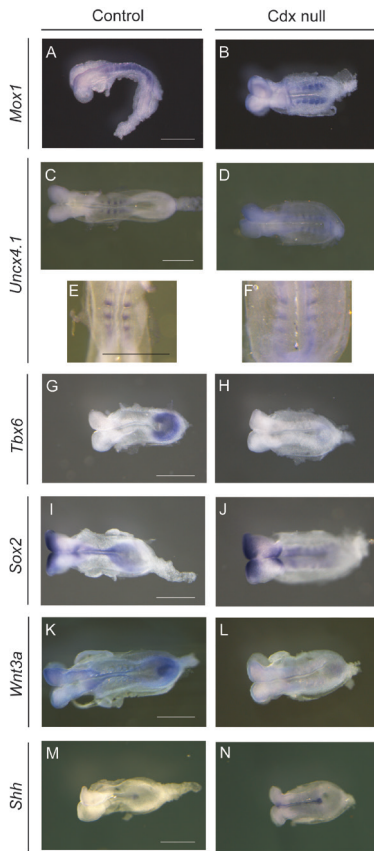


Figure 1. Cdx null mutants make occipital somites and neural tissues but fail to generate more posterior tissues

(A-N) Expression of different markers in E8.5 Cdx null embryos. (A,B) *Mox1* in a control (A) and a Cdx null (B) embryo. (C-F) *Uncx4.1* in a wild type (C and E, 7 somites) and Cdx null mutant (D and F, 5 somites). E and F are close ups of C and D. (G,H) *Tbx6* in a control (G, 6 somites) and Cdx null mutant (H, 5-somites). (I,J) *Sox2* in a control (I, 6 somites) and Cdx null mutant (J, 5 somites). (K,L) *Wnt3a* in a control (K, 9 somites) and Cdx null mutant (L, 5/6 somites). (M,N) *Shh* in a wild type (M, 5 somites) and Cdx null mutant (N, 4 somites). Anterior is to the left in A-D and G-N, and to the top in E,F. Scale bars: 0.5 mm. See also Fig. S1.

2

(one mutant generating a small sixth somite). *Mox1*, a marker of differentiated somites, was expressed in the paraxial mesoderm and confirmed the presence of 5 somites in the Cdx null embryos (Fig. 1A,B). These somites were correctly patterned along their antero-posterior (A-P) axis, as revealed by *Uncx4.1* expression that identifies the posterior somitic compartments (Mansouri et al., 1997) in both mutants and controls (Fig. 1C-F). *Mesp2* normally marking the anterior presomitic mesoderm (PSM) was not expressed in the mutant embryos (Fig. S1G,H). We conclude that Cdx null mutants arrest their axial elongation after the occipital somites have been generated.

Anterior tissues are generated in early gastrulating Cdx null embryos but progenitor cells for trunk and tail fail to be maintained

E8.5 Cdx null embryos and age-matched controls were submitted to in situ hybridization to detect the expression of genes marking recent mesoderm and neurectoderm generated from

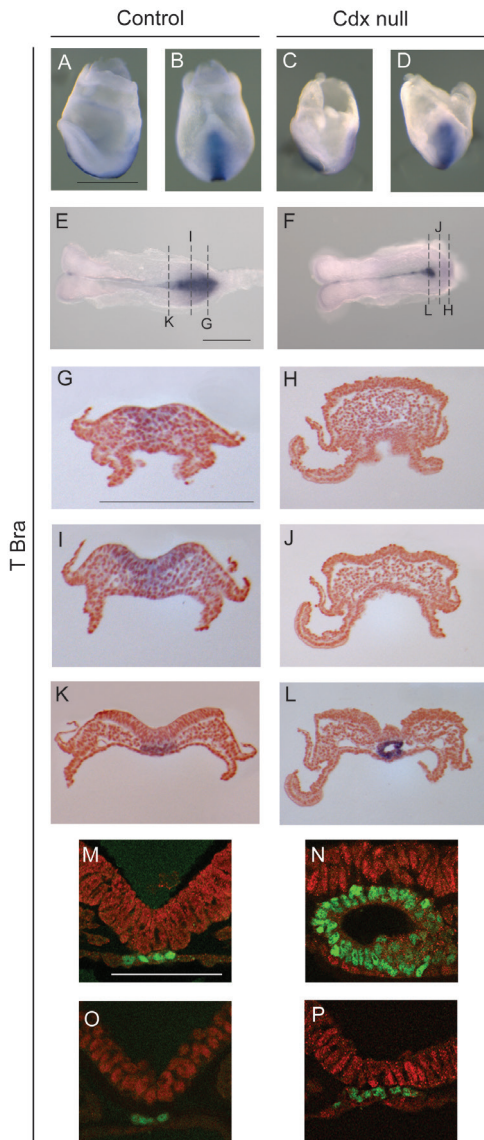


Figure 2. Early generation of nascent mesoderm takes place but ends at the 5-somite stage because of loss of primitive streak activity. (A-L) Expression of *T Brachyury* in E7.5 wild types (A,B, lateral and posterior views) and Cdx null mutant (C, D, lateral and posterior view), and in E8.5 wild type (E, dorsal view, 4 somites) and Cdx null mutant (F, dorsal view, 4 somites). (G-L), transversal sections of E and F. (M-P) Immunofluorescent staining of transversal sections of a E8.5 embryos wild type (M,O) and Cdx null (N,P) at the level of the posteriormost notochord (M,N) and at more anterior levels (O,P). N and P are from the same embryo, whereas M and O are from 2 different control embryos. Anterior is to the left in A,C,E,F; dorsal is up in G-P. Scale bars: 0.5 mm.

the posterior growth zone. *Tbx6*, a marker of the PSM was hardly expressed in the posterior part of Cdx null embryos whereas transcripts were present in the most recently generated paraxial mesoderm in the controls (Fig. 1G,H), suggesting an arrest of mesoderm generation in the mutant. Posterior expression of *Sox2*, marking the neurectoderm was lower in Cdx null embryos, than in controls (Fig. 1I,J). The posterior growth zone of Cdx null mutants thus severely loses its activity in generating nascent mesoderm and neurectoderm. *Wnt3a* expression is also considerably lower in the growth zone of Cdx null mutants versus controls (Fig. 1K,L). The notochord of Cdx null mutants expressed *Shh* (Fig. 1M,N) and *T Brachyury*

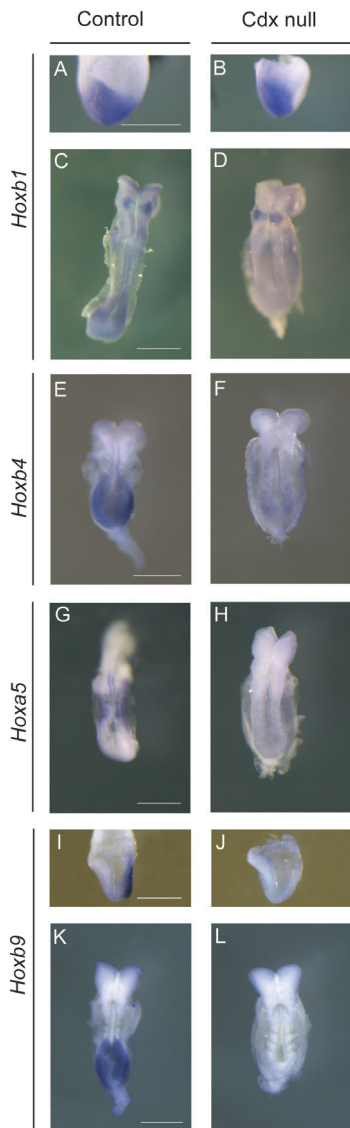


Figure 3. Expression of Hox genes in control and Cdx null embryos. (A-D) Whole mount In situ hybridization with a probe for *Hoxb1* on control (A, C) and Cdx null (B, D) embryos at E7.5 (headfold stage, posterior view, A,B) and at E8.5 (C, 10 somites, and D, 5 somites). (E,F) Expression of *Hoxb4* in E8.5 control (E, 5 somites) and Cdx null embryo (F, 5 somites). (G, H) expression of *Hoxa5* in E8.5 control (G, 9 somites) and Cdx null mutant (H, 5 somites). (I -L) Expression of *Hoxb9* in 1-somite stage control (I) and Cdx null embryos (J), and in E8.5 control (K, 7 somites) and Cdx null embryos (L, 5 somites). Scale bars: 0.5 mm. See also Fig. S2.

2

(Fig. 2E,F). The posterior end of the notochord revealed by the two markers had a widened appearance recognized as tubular by examination of transverse sections, stained for RNA and protein detection of *T Bra* (Fig. 2E,L,N compared to E,K,M). The notochord never expresses Cdx genes and should not be directly affected by the loss of Cdx expression. The tubular end of the notochord is reminiscent of the same feature in mammalian embryos that develop as flat discs [C. Viehbahn, personal communication and (Haldiman and Gier, 1981)], and may thus result from the fact that the Cdx null embryos are much flatter than their controls due to their severe posterior truncation. The notochord of Cdx null mutants at more anterior levels

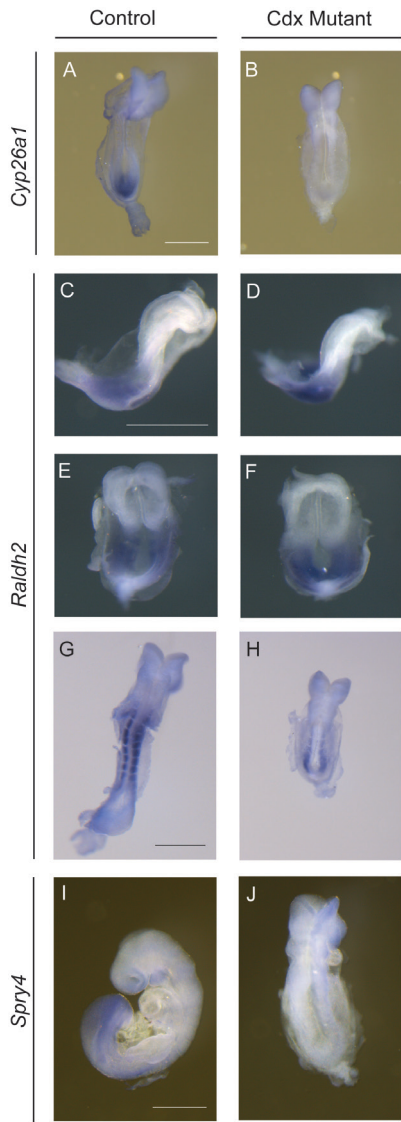


Figure 4. Fgf signaling is lost in the posterior growth zone of Cdx null mutants. (A, B) Expression of *Cyp26a1* in E8.5 wild type (A, 8 somites) and Cdx null mutant embryos (B, 5 somites). (C, F) Expression of *Raldh2* in 2-somite embryos wild type (C, E, lateral and dorsal views, respectively) and Cdx null mutant (D, F, lateral and dorsal views, respectively). (G, H) *Raldh2* expression in E8.5 control (G, 10 somite) and Cdx null (H, 4 somite). (I, J) Expression of the Fgf signaling target *Spry4* in E8.5 wild type (I, 11 somites) and absence of expression in Cdx null mutant (J, 5 somite). Anterior is to the top in A, B, E, F, and G- J, and to the right in C, D. Scale bars: 0.5 mm.

does not show this tubular feature (Fig. 2P). A most striking feature in Cdx null mutants is the absence of *T Brachyury* expression in the primitive streak region posterior to the notochord end at E8.5, whereas the gene is strongly expressed in the streak and in adjacent tissues in age-matched controls (Fig. 2F compared to E). Investigations in earlier, head fold (E7.5) stage embryos revealed that their primitive streak region expresses *T Bra*, and is indistinguishable from that in age-matched controls (Fig. 2A-D). These observations strongly suggest that the progenitors for embryonic axial tissues along the primitive streak at E7.5 normally generate anterior mesoderm in Cdx null mutants, whereas they fail to do so after 5 somites have

been generated. Serial transverse sections through E8.5 embryos hybridized with a *T Bra* probe show that while *T*-positive nascent mesoderm is emerging from the *T*-positive primitive streak in the control, this is not the case in the mutant (Fig. 2H,J versus G,I). This suggests that no new mesoderm has emerged from the inactive primitive streak at E8.5 (Fig. 2H,J).

Anterior Hox genes are well induced in the primitive streak and correctly expressed in anterior tissues whereas more posterior Hox genes are not expressed

The expression of anterior Hox genes was similarly initiated in *Cdx* null mutants and controls. Hox genes are initially transcriptionally induced in the posterior primitive streak at the late mid-streak stage (E7.0/E7.2), and their expression domains spread anterior-wards along the streak and adjacent tissues, in a way that is temporally colinear with the position of the genes in their cluster (Gaunt et al., 1986; Deschamps and Wijgerde, 1993; Forlani et al., 2003; Iimura and Pourquie, 2006). These expression domains then extend further anteriorly in embryonic tissues, eventually reaching gene-specific rostral boundaries. *Hoxb1* is first expressed in the posterior streak at E7.2, and its expression domain has reached the anterior part of the streak by the head fold stage (E7.5) (Forlani et al., 2003) in *Cdx* null mutants like in controls (Fig. 3A, B). At somite stages the anterior expression pattern of *Hoxb1* in the mutant and control is the same, as witnessed by the restricted expression domain at the level of rhombomere 4 (r4) (Fig. 3C,D). However, a reduction of the expression level of the gene in posterior tissues was observed in the mutant, in the primitive streak area reported above to be losing its activity (Fig. 3D). The expression of more posterior Hox genes was analysed in *Cdx* null embryos and controls. *Hoxb4*, the rostral expression domain of which normally reaches the posterior hindbrain and somite 5/6 in the mesoderm (Gould et al., 1998), had an expression boundary caudal to the level of the fifth somite in *Cdx* null embryos, and its expression decayed posteriorly (Fig. 3E,F). E8.5 *Cdx* null embryos did not express *Hoxa5* and *Hoxb8*, normally expressed in trunk tissues (rostral limits in posterior hindbrain and somite 6/7 for *Hoxa5*, and below somite 5 in the neural tube and somite 11 in the paraxial mesoderm for *Hoxb8*) (Larochelle et al., 1999; van den Akker et al., 1999; Young et al., 2009), (Fig. 3G,H and Fig. S2). The same holds true for *Hoxa9* and for *Hoxb9*, two more 5' and later initiated Hox genes that are expressed at trunk levels (Fig. 3K,L and Fig. S2). We conclude that the initial transcription of the first Hox gene of the cluster takes place correctly in the primitive streak of *Cdx* null embryos at early stages. The transcription domain of this 3', early Hox gene normally expands anteriorly together with the emerging tissue that will form the rhombencephalic and occipital structures. The posterior part of the expression of these 3' genes later on fades away as the growth zone becomes inactive at the 5-somite stage, and more 5' (posterior) Hox genes are not expressed.

Fgf is key to *Cdx*-dependent tissue generation from axial progenitors and rescues the posterior truncation in *Cdx2* null mutants

Cdx factors have been suggested to regulate the gene encoding the retinoic acid (RA) degrad-

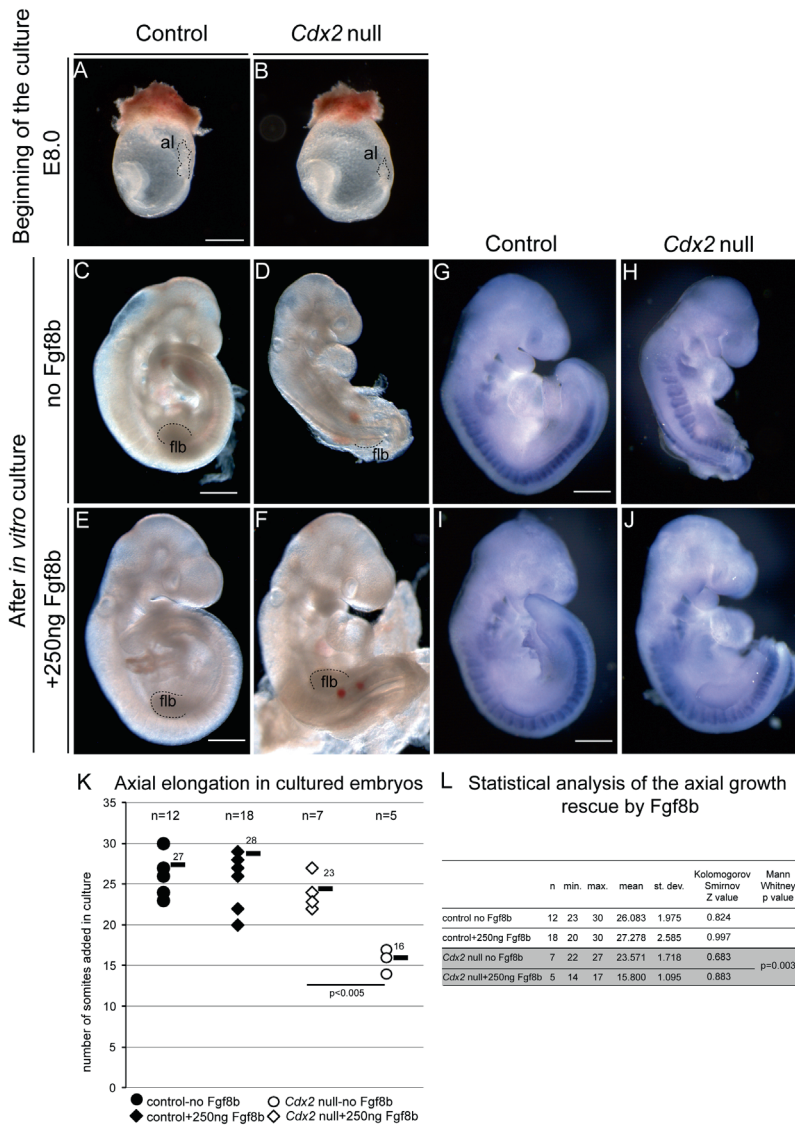


Figure 5. Addition of Fgf8 to the whole embryo culture medium of *Cdx2* null embryos rescues their axial elongation. (A, B) E 8.0 (early somite) embryos at the start of the culture (A, wild type, B, *Cdx2* null). (C –F) Embryos after their culture for the same period without (C, D) or with (E,F) Fgf8 added to the culture medium. C, E are controls, and D, F are *Cdx2* null mutants. (G–J) *Mox1* expression in another set of *Cdx2* null mutants (H,J) and controls (G,I) that have been cultured for the same period with (I,J) and without Fgf8 (G,H). (K) Comparison of posterior elongation of *Cdx2* null and control embryos cultured without and with Fgf8. Y axis, total number of somites generated in culture; median values, bars on the graph; n, number of embryos; several experimental data are superimposed as one symbol in the graph as they had the same value; p, statistical significance (L), Statistical analysis of the axial growth rescue of *Cdx* mutants by Fgf8 using the Mann Whitney test. al, allantois; flb, forelimb bud. Anterior is to the right in A,B and up in C–J. Scale bars: 0.5 mm. See also Fig. S3.

ing enzyme *Cyp26a1* directly and positively (Savory et al., 2009; Young et al., 2009). *Cyp26a1* was not transcribed at all in early somite *Cdx* null mutants (Fig. 4A,B), at a stage when this gene is normally expressed posteriorly and allows the growth zone to clear the RA diffusing from the somites. E8.5 *Cdx* null embryos have stopped generating PSM tissue beyond the last formed somite, and they express the RA-synthesizing enzyme *Raldh2* at high level down to the growth zone, whereas *Raldh2* expression in wild types is restricted to the somites and anterior PSM (Delfini et al., 2005), at a distance from the anterior streak (Fig. 4C-H). *Cdx* null mutants thus, unlike controls, synthesize RA within their growth zone at E8.5. Given the balanced antagonism between the RA and Fgf pathways during posterior embryonic morphogenesis (Diez del Corral and Storey, 2004; Ribes et al., 2009), we set out to test the involvement of the Fgf signaling pathway in causing the posterior axial truncations of *Cdx* null mutants. Fgf signaling activity, revealed by *Spry4* expression (Naiche et al., 2011), was completely lost in the posterior part of E8.5 *Cdx* null mouse embryos (Fig. 4I,J).

We designed whole embryo culture experiments to challenge the critical involvement of Fgf loss in causing the posterior truncation phenotype of *Cdx* mutants. Envisaging rescue attempts on early embryos from the crosses used to generate *Cdx* null embryos was unrealistic given the extremely low yield of these mutants, which cannot be genotyped before the culture. We therefore turned to *Cdx* mutants of the allelic series that are less severely impaired in their development, and easier to generate. *Cdx2* null embryos arrest their development at E10.5, and never have more than 17 somites, no matter whether they are analysed at E9.5 or E10.5 (Chawengsaksophak et al., 1997; Chawengsaksophak et al., 2004; van de Ven et al., 2011), whereas controls typically have about 25 somites at E9.5, and 35 somites at E10.5 [average of many experiments, and see also (Kaufman, 1995)]. We cultured whole E7.5 (presomite)/ E8.0 (early somite) *Cdx2* null mutant and control embryos for the same period of two days *in vitro*, in the presence or in the absence of added recombinant Fgf8 (Fgf8 isoform b). We scored the somite number that these embryos generated during the culture period. Fgf8 exposure was found to significantly rescue the deficit in axial tissue growth of the mutants (Fig. 5F compared to D). *Cdx2* mutant embryos cultured with Fgf8 (n=7) made on average 23 somites during the culture, whereas they only made 16 somites without supplemented Fgf8 (n=5) (Fig. 5K). The rescue of the posterior truncation of *Cdx2* null embryos by Fgf8 was not complete as the PSM remained shorter in the Fgf8-rescued mutants than in controls. Cultured embryos of the four series (controls and mutants cultured with and without Fgf8) were submitted to In situ hybridization with a *Mox1* probe (Fig. 5 G-J) and their somites counted again. This confirmed the rescue of the posterior truncation of the *Cdx2* mutants by Fgf8 (Fig. 5 K). The restoration of posterior axial extension of *Cdx2* mutant embryos by Fgf8 was also documented by measuring the axial length beyond the forelimb buds in stage-matched mutants cultured with (n=4) and without Fgf8 (n=4). The axial portion added to the embryos posterior to the forelimb bud was significantly longer for the mutant embryos cultured with added Fgf8 (median value 472 mm) than for mutants cultured in the same conditions but without Fgf8 (median value 320 mm) (p=0.002) (data not shown). A significant

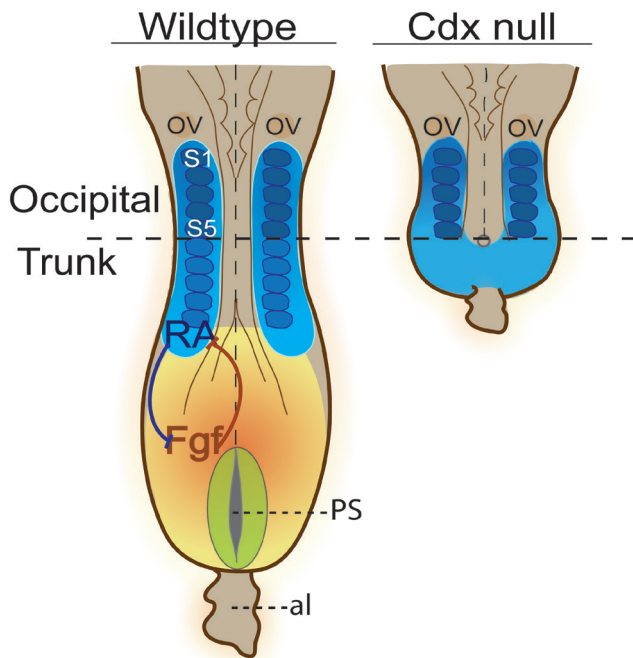


Figure 6. Schematic representation of the loss of Fgf signaling and posterior growth zone in Cdx null mutant embryos. Schematic dorsal view of a E8.5 wild type (lower left) and Cdx null mutant (lower right) embryos. Domains of RA synthesis are in blue and Fgf signaling in orange to yellow; posterior growth zone is in green. PS, primitive streak; al, allantois. Anterior is up in the lower panels.

2

rescue of axial extension is thus taking place when the mutants are grown in the presence of exogenous Fgf8. Importantly, exogenous Fgf8 allowed the *Cdx2* mutant embryos to generate 7 somites more than they ever generate *in vivo* (Fig. 5K). Fgf8-rescued *Cdx2* mutants also re-express *Cyp26a1* at their posterior end similarly to wild types (Fig. S3). The restoration of axial elongation and posterior gene expression by supplemented Fgf, suggests that Fgf signaling reactivates the growth zone in Cdx mutants. We conclude that decreased Fgf signaling in the posterior growth zone in Cdx mutants is crucially involved in causing the exhaustion of tissue emergence from this growth zone. The data are summarized in Fig. 6.

Discussion

Cdx genes are obligatory players in the emergence of the entire trunk and tail. They implement the dichotomy between the pre- and post-occipital tissues. It is known that head tissues are generated early during vertebrate embryogenesis whereas the rest of the axial structures are added subsequently from the posterior growth zone. The absence of active Cdx genes does not affect the generation of head and occipital tissues (the “extended head”), but it prevents trunk and tail tissues to be formed as a result of the depletion of axial progenitor populations from the growth zone. Examination of *T Brachyury* expression in the Cdx null embryos

confirms that early nascent mesoderm emerges normally but stops being generated after 5 somites have formed. The Cdx-dependence of axial elongation is confined to the post-occipital tissues.

These data, together with recent research in lower bilaterians support the hypothesis that the role of Cdx genes in “post head” body extension is ancestral, and exists in arthropods with short germ band development. Cdx/*caudal* must have been involved in generating post-head axial structures since before protostomes and deuterostomes diverged from each other, as witnessed by the obligatory role of Cdx to generate post-head tissues in the short-germ band beetle *Tribolium castaneum*, the crustacean *Artemia franciscana* (Copf et al., 2004), and the intermediate-germ band cricket *Gryllus bimaculatus* (Shinmyo et al., 2005). *Caudal* is therefore an ancestral master organizer of post-head morphogenesis. Its role has been conserved in all animals that sequentially add their trunk and tail structures from a posterior growth zone, and has been reduced in the derived higher dipterans such as *Drosophila melanogaster* (Olesnický et al., 2006).

While *Drosophila* *caudal* does not directly regulate Hox genes, regulatory interactions are known to occur between mouse Cdx and Hox genes [(Young et al., 2009) and references therein]. We show here that Cdx genes are clearly not required for the transcriptional activation of the early (3') genes of the Hox clusters. Transcriptional initiation of 3' Hox genes in the primitive streak is not affected by the absence of Cdx function, and 3' Hox expression in the hindbrain is intact, in agreement with the fact that anterior morphogenesis and signaling are normal in the mutants. In the absence of Cdx activity no axial tissue emerges after occipital somites have been generated, and therefore the later and more 5' Hox genes are not expressed.

Our data in the mouse suggest that Fgf signaling works downstream of Cdx in driving post-occipital tissue emergence. Strikingly, in *Tribolium*, *Tc-Fgf8* is expressed in a region of the posterior growth zone involved in axial elongation, and the Fgf signaling pathway in this insect has been suggested to play a role in posterior mesoderm formation and expansion (Beermann and Schroder, 2008; Beermann et al., 2011). It could therefore well be that both Cdx and its downstream Fgf signaling have been evolutionary conserved in permitting post-head axial extension. The rescuing activity on axial growth of adding Fgf8 during culture of *Cdx2* mutant embryos suggests that FgfR1 signaling is a main contributing intermediate between Cdx and posterior axial extension. The spectrum of posterior abnormalities of Cdx mutants matches well with the phenotype of *FgfR1* mutants. The posterior notochord of Cdx null embryos is thickened, and so is the notochord in *FgfR1* mutant embryos. Chimeric embryos generated with *FgfR1* null ES cells form ectopic neural structures (Deng et al., 1997), and this defect was observed in partial Cdx mutants as well (van de Ven et al., 2011). *FgfR1* mutants are posteriorly truncated and exhibit neural tube closure defects. The neural tube of the severely truncated *FgfR1* null mutants described by Hoch and Soriano (Hoch and Soriano, 2006) remains open along the entire rostro-caudal axis (craniorachischisis), as a consequence of the severe posterior truncation causing a deficit of tension normally facilitating

closure [(Hoch and Soriano, 2006) and references therein]. Cdx null embryos are as severely truncated as the *FgfR1* mutants studied by Hoch and Soriano, suggesting that their lack of neural tube closure may be a consequence of their disrupted FgfR1 signaling.

Cdx null mutants arrest posterior elongation of mesoderm and neurectoderm, and downregulate both Fgf and canonical Wnt signaling in the growth zone. These data support a crucial role of Cdx-dependent Wnt and Fgf signaling in the control of post-occipital axial growth at the level of the maintenance of the bipotent neuro/mesodermal (N/M) progenitors demonstrated in the mouse tailbud (Cambray and Wilson, 2002; Cambray and Wilson, 2007; Tzouanacou et al., 2009).

Material and methods

Mice

All mice were in the C57Bl6j/CBA mixed background. *Cdx2* heterozygotes and *Cdx1* and *Cdx4* null mutant mice as well as the protocols to genotype them have been described previously (Subramanian et al., 1995; Chawengsaksophak and Beck, 1996; van Nes et al., 2006). Generation and genotyping of the strain carrying the *Cdx2* conditional allele was described by (Stringer et al., 2012). Epiblast-specific *Cdx2* null mutants were obtained by crossing *Cdx2* floxed homozygotes and *Cdx2*^{+/+} *Sox2Cre* transgenic mice (Hayashi et al., 2002). Cdx null embryos were generated by crossing *Cdx1*^{-/-} *Cdx2*^{lox/lox} *Cdx4*^{-/-} females with *Cdx1*^{-/-} *Cdx2*^{+/+} males, which carry the *Sox2Cre* transgene. Embryos of the genotype *Cdx1*^{-/-} *Cdx2*^{-/-} *Cdx4*^{+/+} were obtained in the same cross. All experiments using mice were performed in accordance with the institutional and national guidelines and regulations, under control of the “Dutch committee for Animals in Experiments”, and under the licences required in the Netherlands.

Histology, and gene expression analysis

For histological analysis, tissues were fixed with 4% paraformaldehyde (PFA) overnight at 4°C. Whole mount In situ hybridization of mutant and control embryos was performed according to (Young et al., 2009). Embryos were imbedded in plastic (GMA Technovit type 8100) and sectioned at 7 µm. For immunofluorescence staining 70 µm vibratome sections were made from embryos embedded in 4% low melting point agarose. The antibodies were anti-Sox2 (polyclonal rabbit anti-mouse Sox2, Millipore Cat number AB5603), and anti-T Bra (polyclonal goat anti- mouse T, Santa Cruz Cat number SC-17743). Counterstaining was with DAPI.

Whole embryo culture

Embryos were cultured for 48 hours as described by (Bialecka et al., 2010). Each experiment contained control and Cdx mutant embryos with and without Fgf8. At the end of the culture period embryos were fixed in 4% PFA overnight at 4°C, and photographed. Somites were counted using a Leica MZ16FA microscope with a DFC480 camera. Recombinant Fgf8 (iso-

form b) was purchased from R&D systems (Cat number 423-F8).

Statistical analysis

The Mann-Whitney U test was used to analyse the significance of the difference between the number of somites added, and the difference between the length of axial tissue added beyond the forelimb bud of *Cdx* mutant embryos cultured for 2 days with or without Fgf8. The Mann-Whitney U test was chosen because the data sets for each genotype were not normally distributed (Z values obtained from the Kolmogorov-Smirnov test for each genotype were higher than 0.05).

Acknowledgments

We thank E. J. Stringer (Leicester) for generating the *Cdx2* conditional targeting construct in Leicester, and T. Young (Singapore) for generating the targeted mice in Utrecht. We thank M. Reijnen and the animal care staff of the Hubrecht Institute for their help, and Jeroen Korving for assistance with histology. We thank the following colleagues for probes: A. Mansouri (*UncX4.1*), V. Papaioannou (*Tbx6*), D. Stott (*T Brachury*), A. McMahon (*Shh* and *Wnt3a*), Y. Saga (*Mesp2*), S-L Ang (*Mox1*), R. Krumlauf (*Hoxb1*, *Hoxb4* and *Hoxb9*), L. Jeannotte (*Hoxa5*), M. Torres (*Hoxa9*), A. Naiche and M. Lewandoski (*Spry4*), P. Dolle (*Raldh2* and *Cyp26a1*). We acknowledge support from the Dutch Earth and Life Sciences (NWO ALW). This work was also supported by a grant from the Dutch government to the Netherlands Institute for Regenerative Medicine (NIRM, grant No. FES0908). The authors declare no conflict of interests.

References

- Beermann A, Pruhs R, Lutz R, Schroder R. 2011. A context-dependent combination of Wnt receptors controls axis elongation and leg development in a short germ insect. *Development* 138:2793-2805.
- Beermann A, Schroder R. 2008. Sites of Fgf signalling and perception during embryogenesis of the beetle *Tribolium castaneum*. *Dev Genes Evol* 218:153-167.
- Bialecka M, Wilson V, Deschamps J. 2010. Cdx mutant axial progenitor cells are rescued by grafting to a wild type environment. *Dev Biol* 347:228-234.
- Cambray N, Wilson V. 2002. Axial progenitors with extensive potency are localised to the mouse chordoneural hinge. *Development* 129:4855-4866.
- Cambray N, Wilson V. 2007. Two distinct sources for a population of maturing axial progenitors. *Development* 134:2829-2840.
- Chawengsaksophak K, Beck F. 1996. Chromosomal localization of *cdx2*, a murine homologue of the *Drosophila* gene *caudal*, to mouse chromosome 5. *Genomics* 34:270-271.
- Chawengsaksophak K, de Graaff W, Rossant J, Deschamps J, Beck F. 2004. *Cdx2* is essential for axial elongation in mouse development. *Proc Natl Acad Sci U S A* 101:7641-7645.
- Chawengsaksophak K, James R, Hammond VE, Kontgen F, Beck F. 1997. Homeosis and intestinal tumours in *Cdx2* mutant mice. *Nature* 386:84-87.
- Copf T, Schroder R, Averof M. 2004. Ancestral role of caudal genes in axis elongation and segmentation. *Proc Natl Acad Sci U S A* 101:17711-17715.
- Delfini MC, Dubrulle J, Malapert P, Chal J, Pourquie O. 2005. Control of the segmentation process by graded MAPK/ERK activation in the chick embryo. *Proc Natl Acad Sci U S A* 102:11343-11348.
- Deng C, Bedford M, Li C, Xu X, Yang X, Dunmore J, Leder P. 1997. Fibroblast growth factor receptor-1 (FGFR-1) is essential for normal neural tube and limb development. *Dev Biol* 185:42-54.
- Deschamps J, Wijgerde M. 1993. Two phases in the establishment of HOX expression domains. *Dev Biol* 156:473-480.
- Diez del Corral R, Storey KG. 2004. Opposing FGF and retinoid pathways: a signalling switch that controls differentiation and patterning onset in the extending vertebrate body axis. *Bioessays* 26:857-869.
- Forlani S, Lawson KA, Deschamps J. 2003. Acquisition of Hox codes during gastrulation and axial elongation in the mouse embryo. *Development* 130:3807-3819.
- Gaunt SJ, Miller JR, Powell DJ, Duboule D. 1986. Homoeobox gene expression in mouse embryos varies with position by the primitive streak stage. *Nature* 324:662-664.
- Gould A, Itasaki N, Krumlauf R. 1998. Initiation of rhombomeric *Hoxb4* expression requires induction by somites and a retinoid pathway. *Neuron* 21:39-51.
- Haldiman JT, Gier HT. 1981. Bovine notochord origin and development. *Anat Histol Embryol* 10:1-14.
- Hayashi S, Lewis P, Pevny L, McMahon AP. 2002. Efficient gene modulation in mouse epiblast using a *Sox2Cre* transgenic mouse strain. *Mech Dev* 119 Suppl 1:S97-S101.
- Hoch RV, Soriano P. 2006. Context-specific requirements for *Fgfr1* signaling through *Frs2* and *Frs3* during mouse development. *Development* 133:663-673.
- Iimura T, Pourquie O. 2006. Collinear activation of *Hoxb* genes during gastrulation is linked to mesoderm cell ingression. *Nature* 442:568-571.
- Kaufman MH. 1995. *The Atlas of Mouse Development*. London NW1 7DX: Academic Press Limited.
- Kinder SJ, Tsang TE, Quinlan GA, Hadjantonakis AK, Nagy A, Tam PP. 1999. The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo. *Development* 126:4691-4701.
- Larochelle C, Tremblay M, Bernier D, Aubin J, Jeannotte L. 1999. Multiple cis-acting regulatory regions are required for restricted spatio-temporal *Hoxa5* gene expression. *Dev Dyn* 214:127-140.
- Lawson KA, Meneses JJ, Pedersen RA. 1991. Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development* 113:891-911.

- Mansouri A, Yokota Y, Wehr R, Copeland NG, Jenkins NA, Gruss P. 1997. Paired-related murine homeobox gene expressed in the developing sclerotome, kidney, and nervous system. *Dev Dyn* 210:53-65.
- Naïche LA, 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.
- Olesnický EC, Brent AE, Tonnes L, Walker M, Pultz MA, Leaf D, Desplan C. 2006. A caudal mRNA gradient controls posterior development in the wasp *Nasonia*. *Development* 133:3973-3982.
- Ribes V, Le Roux I, Rhinn M, Schuhbaur B, Dolle P. 2009. Early mouse caudal development relies on crosstalk between retinoic acid, Shh and Fgf signalling pathways. *Development* 136:665-676.
- Savory JG, Bouchard N, Pierre V, Rijli FM, De Repentigny Y, Kothary R, Lohnes D. 2009. Cdx2 regulation of posterior development through non-Hox targets. *Development* 136:4099-4110.
- Savory JG, Mansfield M, Rijli FM, Lohnes D. 2011. Cdx mediates neural tube closure through transcriptional regulation of the planar cell polarity gene Ptk7. *Development* 138:1361-1370.
- Shinmyo Y, Mito T, Matsushita T, Sarashina I, Miyawaki K, Ohuchi H, Noji S. 2005. caudal is required for gnathal and thoracic patterning and for posterior elongation in the intermediate-germband cricket *Gryllus bimaculatus*. *Mech Dev* 122:231-239.
- Stringer EJ, Duluc I, Saandi T, Davidson I, Bialecka M, Sato T, Barker N, Clevers H, Pritchard CA, Winton DJ, Wright NA, Freund JN, Deschamps J, Beck F. 2012. Cdx2 determines the fate of postnatal intestinal endoderm. *Development* 139:465-474.
- Subramanian V, Meyer BI, 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.
- Tam PP, Beddington RS. 1987. The formation of mesodermal tissues in the mouse embryo during gastrulation and early organogenesis. *Development* 99:109-126.
- Tzouanacou E, Wegener A, Wymeersch FJ, Wilson V, Nicolas JF. 2009. Redefining the progression of lineage segregations during mammalian embryogenesis by clonal analysis. *Dev Cell* 17:365-376.
- van de Ven C, Bialecka M, Neijts R, Young T, Rowland JE, Stringer EJ, Van Rooijen C, Meijlink F, Novoa A, Freund JN, Mallo M, Beck F, Deschamps J. 2011. Concerted involvement of Cdx/Hox genes and Wnt signaling in morphogenesis of the caudal neural tube and cloacal derivatives from the posterior growth zone. *Development* 138:3451-3462.
- van den Akker E, Forlani S, Chawengsaksophak K, de Graaff W, Beck F, Meyer BI, Deschamps J. 2002. *Cdx1* and *Cdx2* have overlapping functions in anteroposterior patterning and posterior axis elongation. *Development* 129:2181-2193.
- van den Akker E, Reijnen M, Korving J, Brouwer A, Meijlink F, Deschamps J. 1999. Targeted inactivation of *Hoxb8* affects survival of a spinal ganglion and causes aberrant limb reflexes. *Mech Dev* 89:103-114.
- 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.
- Wilson V, Olivera-Martinez I, Storey KG. 2009. Stem cells, signals and vertebrate body axis extension. *Development* 136:1591-1604.
- Young T. 2009. Role of Cdx and Hox genes in posterior axial extension in the mouse. (PhD Thesis) ISBN 9789078675631. Utrecht University.
- Young T, Rowland JE, van de Ven C, Bialecka M, Novoa A, Carapuco M, van Nes J, de Graaff W, Duluc I, Freund JN, Beck F, Mallo M, Deschamps J. 2009. Cdx and Hox genes differentially regulate posterior axial growth in mammalian embryos. *Dev Cell* 17:516-526.



Fig. S1. Phenotype of Cdx null embryos and expression of the nascent somite marker *Mesp2*. (A-F) Control (A,C,E) and Cdx null embryos (B,D,F) at E7.5 (headfold stage; A,B), E8.5 (C, ten somites and D, five somites but similar stage as the control as seen from head developmental aspect), and at E10.5 (E,F). (G,H) Expression of the nascent somite marker *Mesp2* in E8.5 wild type (G) and Cdx null (H) embryos. The wild type is photographed in a manner such that the two *Mesp2* expression stripes are visible (white arrows); they correspond to one of the phases of somitogenesis when one stripe is still visible once the following comes up. The Cdx null embryo did not exhibit any expression of *Mesp2* in the PSM. Scale bars: 0.5 mm.

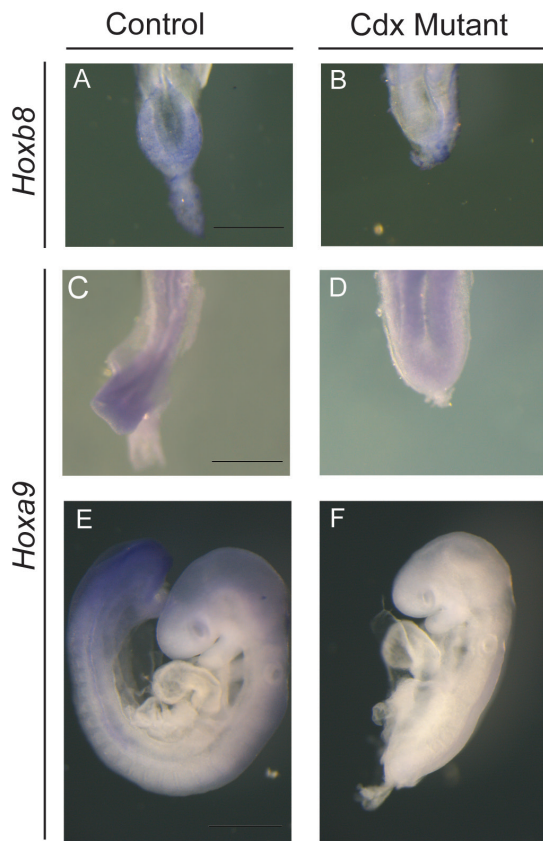


Fig. S2. Expression of *Hoxb8* and *Hoxa9* in *Cdx* null embryos and controls. (A,B) Expression of *Hoxb8* in a E8.5 wild type (A) and *Cdx* null mutant (B) embryos. (C-F) Expression of *Hoxa9* in E8.5 (C,D) and E9.5 (E,F) wild type (C,E) and *Cdx* null (D,F) embryos. Scale bars: 0.5 mm.

2

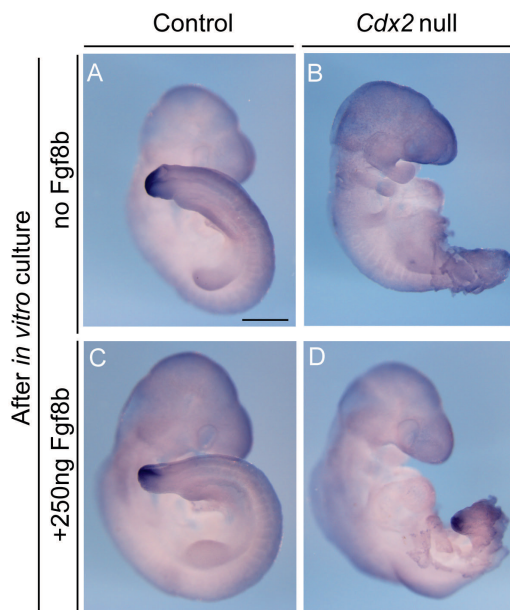


Fig. S3. *Cdx2* null embryos rescued by Fgf8 re-express posterior genes in their growth zone. (A-D) Expression of *Cyp26a1* in *Cdx2* null mutants (B,D) and controls (A,C) that have been cultured for the same 2-day period with (C,D) and without (A,B) Fgf8b. The posterior axis of the *Cdx* mutant embryo cultured with Fgf8 is curled and is therefore longer than it might seem.

2

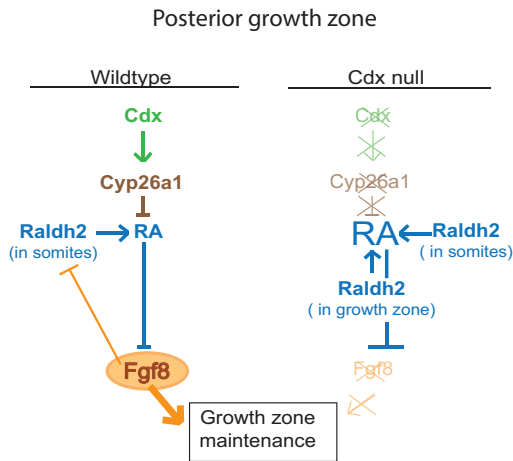


Fig. S4. Schematic of the signaling cascades underlying axial elongation in the posterior growth zone of Cdx null embryos. Genetic interactions downstream of Cdx in the growth zone region of a wild type embryo (left) and a Cdx null mutant embryo (right). These interactions focus on the role of posterior Fgf8 signaling (absent in the mutant). Not mentioned here is the role of Wnt signaling, shown earlier to play a role in axial extension as well in the mouse (Young et al., 2009).

2

Chapter 3

Cdx and T Brachyury co-activate growth signaling in the embryonic axial progenitor niche

Shilu Amin^{1*}, **Roel Neijts**^{1*}, Salvatore Simmini¹, Carina van Rooijen¹, Sander Tan¹, Lennart Kester¹, Alexander van Oudenaarden¹, Menno P. Creyghton¹ and Jacqueline Deschamps¹

**Equal contribution*

¹Hubrecht Institute, Developmental Biology and Stem Cell Research, Uppsalalaan 8, 3584 CT Utrecht, and UMC Utrecht

Amin S*, **Neijts R***, Simmini S, van Rooijen C, Tan S, Kester L, van Oudenaarden A, Creyghton MP and Deschamps J. Cdx and T Brachyury co-activate growth signaling in the embryonic axial progenitor niche,

In press, Cell Reports

**Equal contribution*

Abstract

In vertebrate embryos, anterior tissues are generated early, followed by the other axial structures that emerge sequentially from a posterior growth zone. The genetic network driving posterior axial elongation in mice, and its disturbance in mutants with posterior truncation are not yet fully understood. We show that the combined expression of *Cdx2* and *T Brachyury* is essential to establish the core signature of posterior axial progenitors. *Cdx2* and *T Brachyury* are required for extension of a similar trunk portion of the axis. Simultaneous loss of function of these two genes disrupts axial elongation to a much greater extent than each single mutation alone. We identify and validate common targets for *Cdx2* and *T Brachyury* *in vivo* including *Wnt* and *Fgf* pathway components active in the axial progenitor niche. Our data demonstrate that integration of the *Cdx/Hox* and *T Brachyury* transcriptional networks controls differential axial growth during vertebrate trunk elongation.

Introduction

Our understanding of early post implantation mouse development has increased considerably in recent years, thanks to the refinement of new molecular genetic approaches and the accumulation of morphogenetic information (Rivera-Perez and Hadjantonakis, 2015). This is true in particular for the anterior to posterior growth of embryonic tissues in the three germ layers. The genetic regulation of posterior axial elongation is an evolutionary conserved process in bilaterian animals (Martin and Kimelman, 2009; Neijts et al., 2014). In the mouse, progenitors that supply cells for the different axial tissues of the trunk and tail during the sequential laying down of the anteroposterior structures are present along the primitive streak. Some of these progenitors are generating mesoderm exclusively, whereas a particular population residing in the anteriormost part of the streak represent bipotent neuro-mesodermal precursors (NMPs) that retain the capacity to generate both neural and mesodermal lineages (Cambray and Wilson, 2002, 2007; Tzouanacou et al., 2009; Wymeersch et al., 2016). NMPs have received considerable attention as they possess self-renewing properties (Gouti et al., 2014; Tsakiridis et al., 2014; Turner et al., 2014) and are a key cell population that drives the successive steps of axial tissue generation.

Cdx genes are known to be involved in axial elongation since their inactivation in mice gave rise to embryos with a shortened axis (Chawengsaksophak et al., 1997; van den Akker et al., 2002). All three Cdx genes contribute to this function, the most potent being *Cdx2* (Chawengsaksophak et al., 2004; van Rooijen et al., 2012). Cdx genes are expressed early in the primitive streak and later in the tailbud, where they are required for growth of posterior embryonic tissues until the axis is fully extended. Mutants totally missing active Cdx genes develop anterior structures normally but that they fail to generate any post-occipital tissue (van Rooijen et al., 2012). *Cdx2* null mutants are impaired in generating their axis posterior from somite 7 to 12. Inactivation of *Cdx1* and/or *Cdx4* does not alter axial elongation, but the truncation phenotype of *Cdx2* null mutants is more severe in combination with the inactivation of one of the other two Cdx genes. Rescue experiments *in vivo* (Young et al., 2009) and in embryos in culture (van Rooijen et al., 2012) have indicated that Cdx genes act in axial elongation at least in part by maintaining Wnt and Fgf signaling active in the posterior growth zone suggesting that these may represent key downstream targets for Cdx transcriptional activity.

In addition to Cdx, another transcription factor required for complete posterior axis elongation is T Brachyury (T Bra). *T Bra* is expressed in the primitive streak and early mesoderm at gastrulation stages, and in the growth zone of the tailbud subsequently until around E14.5 (Wilkinson et al., 1990). Similarly to *Cdx2* null mutants, *T Bra* null embryos generate about 7 somites after which axis elongation is impaired. In addition, the neural tube is kinked in its posterior portion and abnormal somites are observed. This is likely due to the fact that *T Bra* plays a role in mesoderm specification as well as in somitogenesis (Martin, 2016; Wilson and Beddington, 1997). We now show that the null mutations in *Cdx2* and *T Bra* synergize in their effects on embryonic axial elongation. We aimed at investigating whether this syner-

gism, added to the similarity between the posterior truncation phenotypes of *Cdx2* and *T Bra* mutants, results from the same molecular mechanism of action. Similarly to the situation for *Cdx2*, *T Bra* regulates the Wnt signaling pathway (Martin and Kimelman, 2008) and the Fgf pathway, as *Fgf8* was identified as a target of *T Bra* in differentiating mouse ES cells (Lolas et al., 2014).

Recent work demonstrated that the co-expression of *T Bra* and the stem cell marker *Sox2* is linked to a core transcriptional signature in these NMPs (Olivera-Martinez et al., 2012; Wymeersch et al., 2016). *T Bra* is required for the activity of the axial progenitors contributing descendants at axial levels posterior to somite 7, the region described above to be dependent on *Cdx2* for its generation. In order to understand whether *Cdx2* like *T Bra* maintains the axial progenitor niche proficient, and how this is executed by these transcription factors, we set out to determine and compare the downstream programs of both *Cdx2* and *T Bra* during posterior axis elongation. We used a homogenous *ex vivo* system based on pre-gastrulation embryo-derived epiblast stem cells (EpiSCs), as a model for the posterior growth zone of the embryo. When induced with Wnt and Fgf, these cells closely represent primitive streak epiblast, including the NMPs described above (Tsakiridis et al., 2014; Tsakiridis and Wilson, 2015). We performed genome-wide binding analysis for direct targets of *Cdx2* and *T Bra* in these cells by chromatin immunoprecipitation (ChIP-seq). We identified an overlapping set of target genes, including members of the Fgf and Wnt signaling cascades. We validated the *Cdx/T Bra* binding regions as transcriptional enhancers of the target genes using *lacZ* reporter assays. We propose that *Cdx2* expression participates in the core signature of posterior NMP progenitors and conclude that *Cdx2* and *T Bra* stimulate axial extension by directly co-activating the Wnt and Fgf growth signaling cascades, both at the level of the axial progenitors themselves and at the level of their niche.

Results

Epiblast stem cells are a valid model of posterior embryonic elongation

Posterior elongation of the axis to generate the trunk and tail tissues occurs from cell progenitors in the primitive streak region in the posterior part of the mouse embryo between the late streak-early somite stages and around E14.5. Several transcription factors and signaling pathways are known to be instrumental in this process, as shown by the posterior truncation phenotype resulting from their invalidation in mutants. This is the case for *T Bra* (Herrmann et al., 1990), *Cdx2* (Chawengsaksophak et al., 2004) and genes of the Wnt (Galceran et al., 1999; Takada et al., 1994) and Fgf (Hoch and Soriano, 2006; Naiche et al., 2011) pathways.

To study these processes we used a model for posterior axis elongation. EpiSCs can be directed towards a primitive streak-like fate by Wnt3a (or the Wnt agonist CHIR99021, Chiron) and Fgf. A proportion of these cells qualify as NMPs that contribute descendants to the elongating trunk and tail tissues (Tsakiridis et al., 2014) (Figure 1A). Induced EpiSCs exhibit specific features of axial stem cells, as shown by the increased availability of the *Sox2*

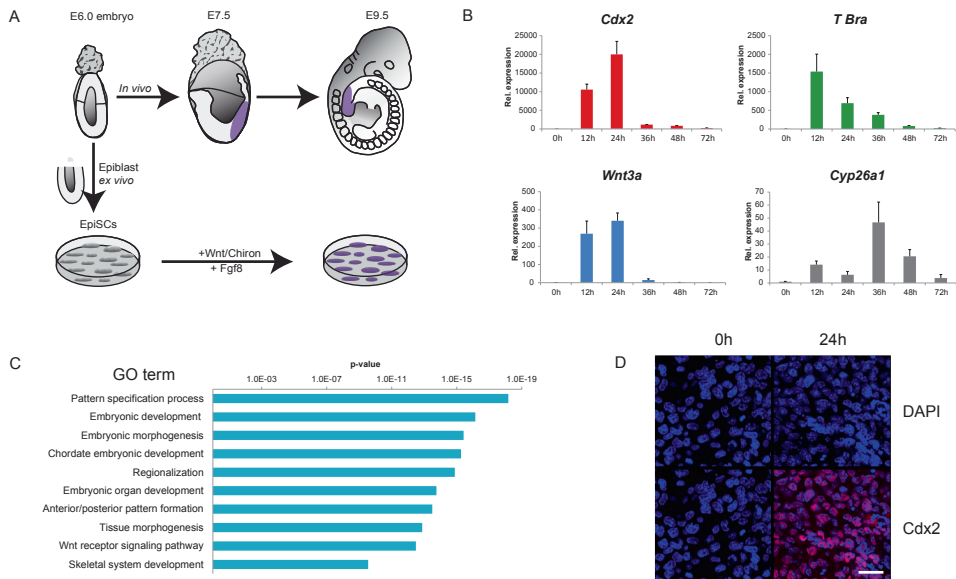


Figure 1. *Cdx2* expression is induced by Wnt and Fgf signaling in EpiSCs. A) Induced EpiSCs as a model system for posterior embryonic development. Posterior gene expression is highlighted in purple. B) Induction time course of *Cdx2* expression and expression of markers for posterior axial extension upon Chiron stimulation for up to 72 hours. Values represent expression relative to 0 hours and normalized to the housekeeping gene *Ppia*. Error bars represent SD of at least two biological replicates. C) Top GO functional categories identified by DAVID analysis of RNA-seq upregulated genes in WT EpiSCs induced for 24 hours by Chiron/Fgf8. The length of the bars corresponds to p-value (x-axis). D) Immunofluorescence staining using anti-*Cdx2* antibody (red) in WT EpiSCs uninduced (0 hours) and 24 hours induced with Chiron/Fgf8. h, hours. Nuclear staining is blue (DAPI). Scale bar is 25 μ m. See also Figure S1.

N1 enhancer, which represents a hallmark of the bipotent NMPs (Takemoto et al., 2011) (Figure S1A). We first measured changes in gene expression in EpiSCs after induction by Wnt3a or the Wnt agonist Chiron and Fgf, following a timecourse up to 72 hours. Typical markers for posterior axial extension, *Cdx2*, *T Bra*, *Wnt3a* and *Cyp26a1*, were transcriptionally highly stimulated in comparison with their expression in non-treated EpiSCs (Figure 1B and Figure S1B), showing that EpiSCs exposed to Wnt and Fgf activate the pathways utilized by progenitors of axial tissues *in vivo*. Validation of EpiSCs as a reliable model for the elongating embryonic axis was further strengthened by the comparison of whole transcriptome analysis of EpiSCs before and after Chiron and Fgf8 stimulation. RNA sequencing (RNA-seq) analysis uncovered 655 genes that were differentially expressed (FDR < 0.05) (Table S1). Gene ontology (GO) analysis of significantly upregulated genes (fold change >2) revealed that the gene families that were affected were predominantly involved in pattern-specification processes and antero-posterior (AP) pattern formation (Figure 1C). *Cdx2* was one of the genes that was

highly induced in the Chiron and Fgf8-treated EpiSCs, and the protein showed homogeneous expression throughout the cell population (Figure 1D).

Direct targets of Cdx2 in Wnt and Fgf- induced EpiSCs and embryo tailbuds

In order to identify the direct targets of Cdx2, ChIP-seq was performed using an anti-Cdx2 antibody in EpiSCs induced with Chiron and Fgf8 for 24 hours. 3682 Cdx2 binding regions were identified from two replicates by MACS (Zhang et al., 2008) (Table S2). By performing motif analysis on regions 200 bp around the summit of peaks, the Cdx2 binding consensus sequence was found to be the top enriched motif (Figure 2A and Figure S2E). The majority of binding regions were localized distal to transcription start sites (Figure S2A). Gene ontology (GO) analysis of Cdx2 ChIP-seq binding regions (fold enrichment > 5) using the 'GREAT basal plus extension rule' (McLean et al., 2010) demonstrated enrichment of genes involved in processes associated to regionalization, AP patterning, and stem cell differentiation and development. Interestingly the only signaling pathway enriched term in the GO analysis was the Wnt signaling pathway (Figure 2B). Wnt pathway genes with active Cdx2 binding regions in their vicinity, as demonstrated by H3K27ac enrichment, are *Fzd10*, *Lef1* and *Wnt5a* (Figure 2C). We assigned the 3682 Cdx2 binding regions to a total of 3970 genes using GREAT (Table S6).

To independently validate these Cdx2 targets in embryos *in vivo*, we also performed RNA-seq analysis in dissected tailbuds of E8.0 Cdx mutant and WT embryos. We compared the transcriptome of 2 to 5-somite aged Cdx triple null embryos, which exhibit posterior truncation, with the transcriptome of *Cdx1-Cdx4* double mutants (*Cdx1-4* null) which are not impaired in their axial elongation, and that of age-matched wild types. We found that the sets of genes deregulated in Cdx null mutants versus WT, and in Cdx null mutants versus *Cdx1-4* null were similar, whereas the comparison between *Cdx1-4* null and WT embryos uncovered only a few genes with a significant expression changes (Figure 2E and Table S3). 172 genes were downregulated and 215 genes were upregulated in Cdx triple null compared to WT embryos (fold change > 1.3 and p-value < 0.05). This confirmed at the gene expression level that Cdx2 is the key player in the process of axial elongation.

Next we determined the overlap between the set of Cdx2 bound loci revealed by the ChIP-seq experiments in WT EpiSCs with the up and downregulated genes in the RNA-seq performed on Cdx triple mutant versus WT embryos (Figure 2D). 43.0% of the genes downregulated, and 27.4% of the genes upregulated in Cdx triple null mutant embryos, had at least one Cdx2-bound region assigned to them suggesting that Cdx2 binding plays a more frequent direct role in gene activation. Most interestingly, several genes of the Wnt and Fgf pathways were bound by Cdx2 and downregulated in Cdx mutants (Figure 2E), convincingly demonstrating that Cdx2 directly stimulates these signaling pathways and that this stimulation is an essential and limiting step for embryonic posterior axial elongation.

A particular category of Cdx2 targets that is worth mentioning in the context of axial elongation concerns the Hox gene clusters. Cdx2 binds to the Hox1 – Hox9 subset of

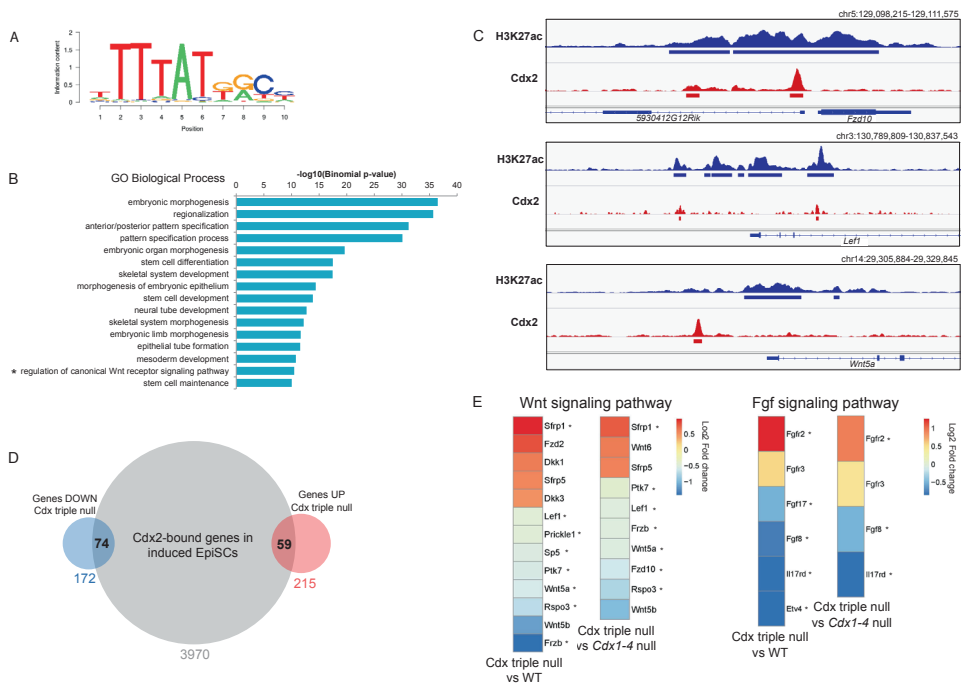


Figure 2. Cdx2 directly targets genes in the Wnt and Fgf signaling pathways. A) Sequence logo of the top enriched motif in Cdx2-bound 200 bp summit regions identified in induced EpiSCs by SeqPos motif tool in Galaxy Cistrome. B) Top overrepresented 'Biological Process' categories identified by GREAT analysis of top significant Cdx2 ChIP-seq binding regions. The length of the bars corresponds to the binomial raw (uncorrected) p-value (x-axis). C) ChIP-seq tracks for H3K27ac (blue) and Cdx2 (red) corresponding to the genomic regions containing *Fzd10*, *Lef1* and *Wnt5a*. Solid bars under each track represent the MACS peak calling identified regions. D) Overlap of genes identified by 'basal plus extension association rule' from Cdx2 ChIP-seq in induced WT EpiSCs (grey circle) with genes upregulated (red circle) and downregulated (blue circle) from Cdx triple null vs WT embryos in RNA-seq analyses. E) Differentially expressed genes in Cdx triple null vs WT and in Cdx triple null vs *Cdx1-4* null embryos linked to Wnt and Fgf signaling pathways with their corresponding Log₂ fold changes. Asterisks indicate genes bound by Cdx2 from ChIP-seq analysis in induced WT EpiSCs. See also Figure S2.

Hox genes in each cluster (Figure S2B, C) and these genes are upregulated in Wnt and Fgf induced EpiSCs and downregulated in Cdx triple null embryos with the exception of *Hoxa1* (Figure S2D). This is in line with the collaborative role of central Hox genes in axial elongation reported previously (Young et al., 2009). Given our previous genetic data on the antagonistic role of Hox13 genes on the Cdx/Wnt/Fgf-supported axial elongation process at the trunk to tail transition in the mouse (van Rooijen et al., 2012; Young et al., 2009), and the inhibitory effect of Hox13 proteins on Wnt signaling documented in transient electroporation studies in chick embryos (Denans et al., 2015), we asked whether the direct targets of Cdx2 in the Wnt and Fgf pathways are also bound by Hox13 gene products. Binding motif analysis for these

Hox gene family transcription factors revealed that Hox13 proteins bind the same consensus sequence as Cdx proteins (Figure S2F). We previously demonstrated that precocious expression of Hox13 proteins, using the *Cdx2* promoter (*Cdx2P*), causes posterior truncation of the embryonic axis (Young et al., 2009). *Cdx2P-Hoxb13* homozygous mice manifested a moderate truncation of their tail. We used transgenic embryos from these mice to investigate whether precocious *Hoxb13* could exert its growth antagonistic effect by binding to the Wnt and Fgf targets of Cdx2. We made use of the FLAG tag in front of the Hoxb13 N-terminus (Figure S2G) to immunoprecipitate chromatin of dissected tailbud versus anterior trunk tissues of *Cdx2P-Hoxb13* embryos, and measured the enrichment of recovered DNA by qPCR for 4 different Cdx2 targets belonging to the Wnt and Fgf pathways (Figure S2H). This revealed that Hoxb13 binds these Cdx2 targets very efficiently *in vivo*. In addition, Cdx2 targets such as *T Bra* – a Cdx2 target essential for axial growth – is strongly downregulated by overexpression of *Hoxc13* in the posterior part of E10.5 transgenic embryos *in vivo* (Young et al., 2009). We conclude that the slowing down of axial elongation by Hox13 proteins may be executed by their direct binding to the same regulatory elements as Cdx2, and thereby arresting Cdx-dependent growth signaling.

Functional validation of Cdx2 direct targets involved in embryonic posterior axial elongation

To verify the involvement of a number of Cdx2-bound loci in the process of axial elongation in embryos, we determined whether the expression of these target genes was affected in Cdx triple null mutant embryos using *in situ* hybridization (ISH). Expression of *Fzd10*, *Wnt5a* and *Fgf8* was decreased in the posterior parts of E8.0 mutant embryos compared to their age-matched controls (white arrows in Figure 3A and S3A). Expression of Cdx2-bound genes was also quantified by qPCR in E7.5 WT and Cdx triple null embryos, as well as in non-induced and induced WT and Cdx triple null EpiSCs, demonstrating that these genes are dependent on Cdx for their transcription (Figure 3A and Figure S3D).

Cdx2 binding regions were tested for enhancer activity using *lacZ* reporter assays in transgenic mice. The Cdx2 binding region upstream of *Fzd10* drives *lacZ* expression specifically in the posterior part of E9.5 transgenic embryos (Figure 3B). Furthermore, the Cdx2 binding regions identified near *Fgf8*, *Wnt5a* and *Spry4-Fgf1* correspond to conserved genomic regions between human and mouse that have been validated previously as enhancers with posterior embryonic activity (Fig S3B, VISTA enhancer browser hs511, hs1472, and hs1640) (Visel et al., 2007).

DNA accessibility analysis (ATAC-seq) showed that the Cdx2 binding regions near *Fgf8*, *Wnt5a*, *Spry4-Fgf1* and *Rspo3*, and to a lesser extent *Fzd10*, are more accessible after induction of WT EpiSCs with Chiron and Fgf (Figure 3B and Figure S3B-D) (peaks distal to *Rspo3* were confirmed to be associated to the *Rspo3* promoter, data not shown). Similar results were found for a Cdx2 target site in the *Cdx2* regulatory region (Benahmed et al., 2008) and a region near *Lef1*, the gene encoding an executive transcription factor of the Wnt pathway

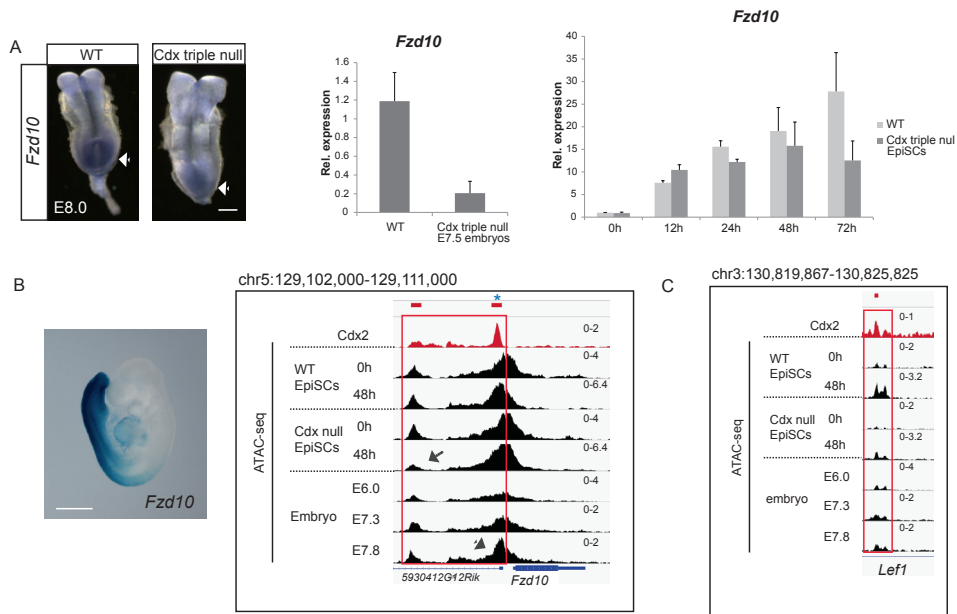


Figure 3. Cdx2 directly acts on chromatin to target gene transcription involved in embryonic posterior axial elongation. A) Left, ISH on E8.0 WT and Cdx triple null embryos using a *Fzd10* probe. White arrows indicate posterior expression. Scale bar is 100 μ m. Right, *Fzd10* expression levels in WT and Cdx triple null E7.5 embryos. Values represent expression relative to internal control gene. Induction time course of *Fzd10* expression in WT and Cdx triple null EpiSCs upon Chiron stimulation for up to 72 hours. Values represent expression relative to WT EpiSCs at 0 hours and normalized to *Ppia*. Error bars represent SD of at least two biological replicates. B) Left, activity of Cdx2-bound *Fzd10* region coupled to a *lacZ* reporter in E9.5 embryos. Scale bar is 1mm. Right, ATAC-seq profiles at the promoter region of *Fzd10* (red box) in uninduced and induced WT and Cdx triple null EpiSCs, and in early embryos. The asterisk highlights the Cdx2 binding region tested in the *lacZ* reporter assay. The arrow shows less opening of the Cdx2 binding regions in induced Cdx triple null EpiSCs. The arrowhead shows this Cdx2 binding region is fully open in embryos when *Cdx2* expression is initiated. C) ATAC-seq profiles at *Lef1* locus. h, hours. Red solid bars (above) represent the MACS peak calling identified regions See also Figure S3.

(Figure S3C and Figure 3C). These Cdx2-bound regions become accessible in induced WT EpiSCs whereas they do not become accessible in induced Cdx null EpiSCs. Moreover, most of these regions only become accessible in E7.8 embryos *in vivo*, consistent with the unavailability of Cdx proteins at earlier stages (Figure 3B, C and Figure S3B-D).

A more comprehensive analysis of Cdx2-dependent enhancers revealed that a large number of genes in the Wnt and Fgf pathways are bound by Cdx2 and also fail to become accessible in induced Cdx triple null EpiSCs (Table S4). Several of these genes also become accessible in E7.8 embryos (highlighted in Table S4). GO analysis revealed regions bound by Cdx2 that are inaccessible in Cdx triple null EpiSCs are enriched for genes associated to transcriptional regulation and include transcription factors that are essential for pattern specifica-

tion (Table S4). This data suggests that *Cdx2* could act as a pioneer transcription factor (Zaret and Mango, 2016) that initiates the expression of important downstream target genes among which Wnt and Fgf signaling components.

Next we determined the overlap between the set of regions that become accessible at E7.8 and those that fail to become accessible in induced *Cdx* null EpiSCs (Figure S3F). 79% (107 out of 136) of these enhancers do bind *Cdx2*, and become accessible in a *Cdx*-dependent way in EpiSCs and in embryos. This high percentage is also in line with a role of *Cdx2* as pioneer factor.

Opening of the chromatin at the *Cdx2* binding enhancers, a prerequisite for target gene transcription in induced EpiSCs and in embryos, therefore depends in most cases on the presence of the *Cdx* protein. Collectively these data indicate that *Cdx2* binds to and activates enhancers of genes belonging to the Wnt and Fgf pathways, and that binding and activation are abolished in *Cdx* triple null EpiSCs and embryos.

***Cdx2* and *T Bra* are co-expressed in the posterior axial progenitor region and double mutants exhibit a truncated phenotype more severe than each single mutant**

Inactivation of *Cdx2* and *T Bra* leads to a posterior truncation of the embryonic axis at a similar axial level. We therefore set out to compare the expression pattern of these two genes in detail. Expression of *Cdx2* in the embryo proper begins in the posterior part of the primitive streak at the late streak stage (E7.2). Expression then spreads rostrally along the streak in epiblast and more weakly in the nascent mesoderm. Transcription of *Cdx2* is strong in the streak region and in the presomitic mesoderm (PSM) at E8.5 and E9.5 (Figure S4A), and fades away by E12.5 (Young et al., 2009).

Initial expression of *T Bra* in the posterior part of the E6.0 egg cylinder precedes primitive streak appearance (Rivera-Perez and Magnuson, 2005). *T Bra* is expressed in the epiblast abutting the streak and in the nascent mesoderm ingressing through the streak at E8.5 and E9.5 (Figure S4A). *T Bra* transcription is downregulated in the tailbud at the end of axial extension around E14.5 (Cambray and Wilson, 2007). Both *Cdx2* and *T Bra* are expressed at high levels in posterior embryonic tissues during the developmental period between E7.5 and E10.5 that corresponds to the generation of trunk tissues (Figure 4A). Both genes start to be downregulated around the trunk to tail transition, resulting in a drop of transcription around E12.5 for *Cdx2* and E14.5 for *T Bra*.

Examination of the distribution of the active chromatin mark H3K27ac in the neighbourhood of *Cdx2* and *T Bra* confirms that these loci are active in E9.5 embryonic tailbud tissues whereas they are not in anterior tissues of the same embryos (Figure 4B). Moreover a similar H3K27ac profile is observed in induced EpiSCs (Figure 4B), confirming that these EpiSCs are a valid model for the posteriorly elongating embryonic tissues.

Cdx2 null mutant embryos are arrested after 7 to 12 somites (van de Ven et al., 2011). *T Bra* null embryos do not generate more than about 7 somites and their neural tube at pos-

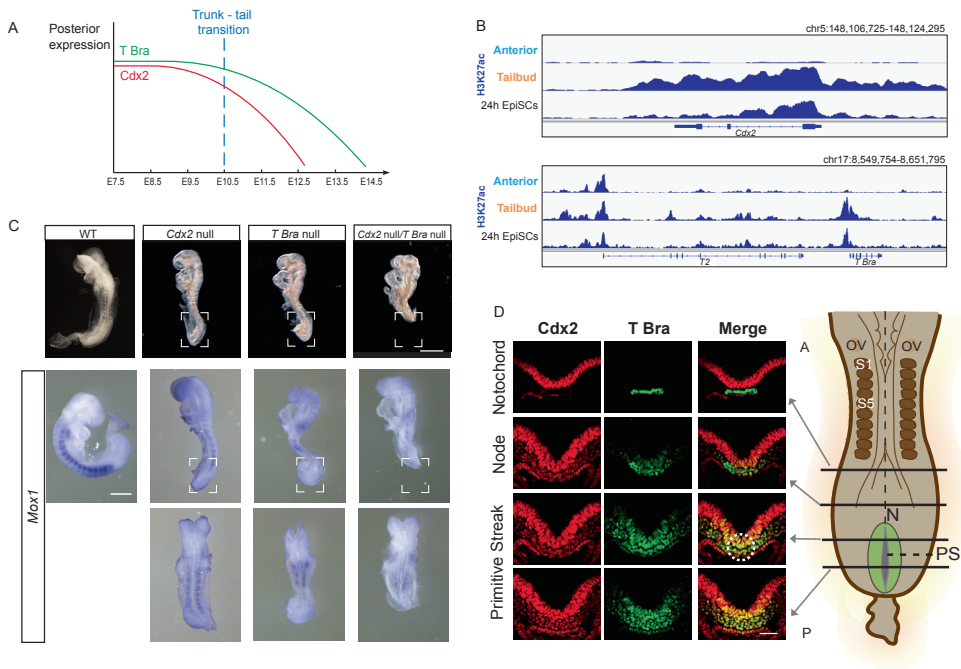


Figure 4. *Cdx2* and *T Bra* are co-expressed in the axial progenitor region and collaborate in driving axial extension. A) Schematic representation of the expression level of *Cdx2* (red) and *T Bra* (green) in the posterior part of developing embryos. Developmental stages are along the x-axis; the trunk to tail transition is indicated (blue dotted line). B) H3K27ac ChIP-seq tracks corresponding to the genomic regions containing *Cdx2* (top panels) and *T2/T Bra* (bottom panels) in anterior tissue, tailbud tissue and WT EpiSCs induced for 24h with Chiron/Egf8. C) Top panel from left to right, WT, *Cdx2* null, *T Bra* null, *Cdx2* null/*T Bra* null E8.5 embryos. Anterior is to the top. The WT embryo has 12 somites, the *Cdx2* null embryo shown here has 10 somites, the *T Bra* null embryo 6 recognizable somites and the *Cdx2* null/*T Bra* null mutant embryo 4-5 identifiable somites. Scale bar is 200 μ m. Bottom panels, expression of somite marker *Mox1* in WT and mutant embryos by ISH. The WT embryo has 15 somites, the *Cdx2* null embryo shown here has 7-8 somites, the *T Bra* null embryo 5 recognizable somites and the *Cdx2* null/*T Bra* null mutant embryo 4 identifiable somites. White brackets highlight tissue missing in *Cdx2* null/*T Bra* null mutant embryos. Scale bar is 100 μ m. D) Left, *Cdx2* and *T Bra* immunofluorescence staining on transversal sections of the posterior region of wild type mouse embryo at E8.0. The axial progenitors reside at the level between the anterior primitive streak and the node (NSB, white dotted circle). Scale bar is 50 μ m. Right, graphic of posterior part of embryo, black solid lines represent level of each section and the posterior growth zone is highlighted in green. A, anterior; P, posterior; N, node; OV, otic vesicle; S, somite PS, primitive streak. See also Figure S4.

terior levels is kinked and abnormal (Rashbass et al., 1994) (Figure S4B). Double mutant embryos were generated from intercrosses of *Cdx2* conditional homozygotes carrying a null allele of *T Bra* and a *Rosa26-Cre ER(T2)* allele. These embryos clearly exhibit a more severely truncated phenotype posteriorly than each single mutant (Figure 4C). Furthermore double mutants miss all posterior tissues that form a tailbud in the WT and single mutants (highlighted in Figure 4C). In the mesoderm that is generated, only 3 to 5 disorganized somites

can be discerned after ISH with the somitic marker *Mox1* (Figure 4C, lower panels). The *T Bra* null mutation affects somite morphogenesis, as visible in *T Bra* null, and *Cdx2* null/*T Bra* null embryos. Embryos lacking both functional *Cdx2* and *T Bra* generate head and occipital structures exclusively. *Cdx2* and *T Bra* thus cooperate in their action in such a way that missing both genes together is much more deleterious for growth of the embryonic trunk than missing each one at a time.

In order to further understand the basis of the *Cdx2/T Bra* double mutant phenotype, we examined the distribution of the *Cdx2* and *T Bra* proteins in the posterior part of the embryo from which the axis extends. Co-staining of transverse sections of the posterior part of E8.5 embryos with antibodies against *T Bra* and *Cdx2* demonstrates that both proteins are present at the same location where NMPs are known to reside (Figure 4D). *Cdx* mutants like *T Bra* mutants form ectopic neural structures at posterior levels (van de Ven et al., 2011; Yamaguchi et al., 1999) (Figure S4D), strengthening the notion that both of these genes control the NMP-dependent growth of the posterior embryonic axis. Mutants in the niche factor *Wnt3a* also exhibit posterior ectopic neural structures (Yoshikawa et al., 1997), indicating that these transcription factors and signaling pathways act in the same network to orchestrate axial extension by modulating the NMP population.

3

***Cdx2* and *T Bra* collaborate in directly activating a Wnt and Fgf gene regulatory network**

Previous work has shown that mouse *Cdx* (Bialecka et al., 2010; van de Ven et al., 2011) and zebrafish *T Bra* (Martin and Kimelman, 2008, 2010) affect axial elongation at least in part by impairing the posterior progenitor niche. We find that both *T Bra* and *Cdx2* in the mouse are co-expressed in the NMP region from which axial tissue expands from throughout the duration of trunk axial elongation. Our observations that *Cdx2* null and *T Bra* null mutations add their effects in impairing axial elongation of the embryonic trunk prompted us to test whether *Cdx2* and *T Bra* stimulate posterior growth signaling by co-activating genes of the Wnt and Fgf pathway.

Cdx2 and *T Bra* are strongly transcribed in the posterior tissues of E7.5 till E10.5 embryos (Figure 4A and Figure S4A). Several genes of the Wnt and Fgf pathways among which *Wnt3a* and *Fgf8*, are also expressed in posterior embryonic tissues, similarly to *Cdx2* and *T Bra* (Figure S5A). We therefore performed ChIP-seq for *T Bra* in induced EpiSCs to investigate whether there are common targets for *T Bra* and *Cdx2* in embryonic posterior tissues. We identified 1215 *T Bra* binding regions from two replicates by MACS (Zhang et al., 2008) (Table S5). The *T Bra* binding motif (Kispert and Herrmann, 1993) is one of the top enriched motifs in the uncovered binding regions, 200 bp around the summit of the peaks (Figure 5A). The other motifs correspond to the binding sequences of other T-box transcription factor-encoding genes (Figure S5C). GO analysis of *T Bra* ChIP-seq data showed enrichment for genes expressed in mesoderm and primitive streak (Figure 5B) with peaks distal to transcription start sites (TSS) (Figure S5B). When comparing the series of *T Bra* and *Cdx2* bound regions, a number of common target genes were identified (Figure 5C). Importantly,

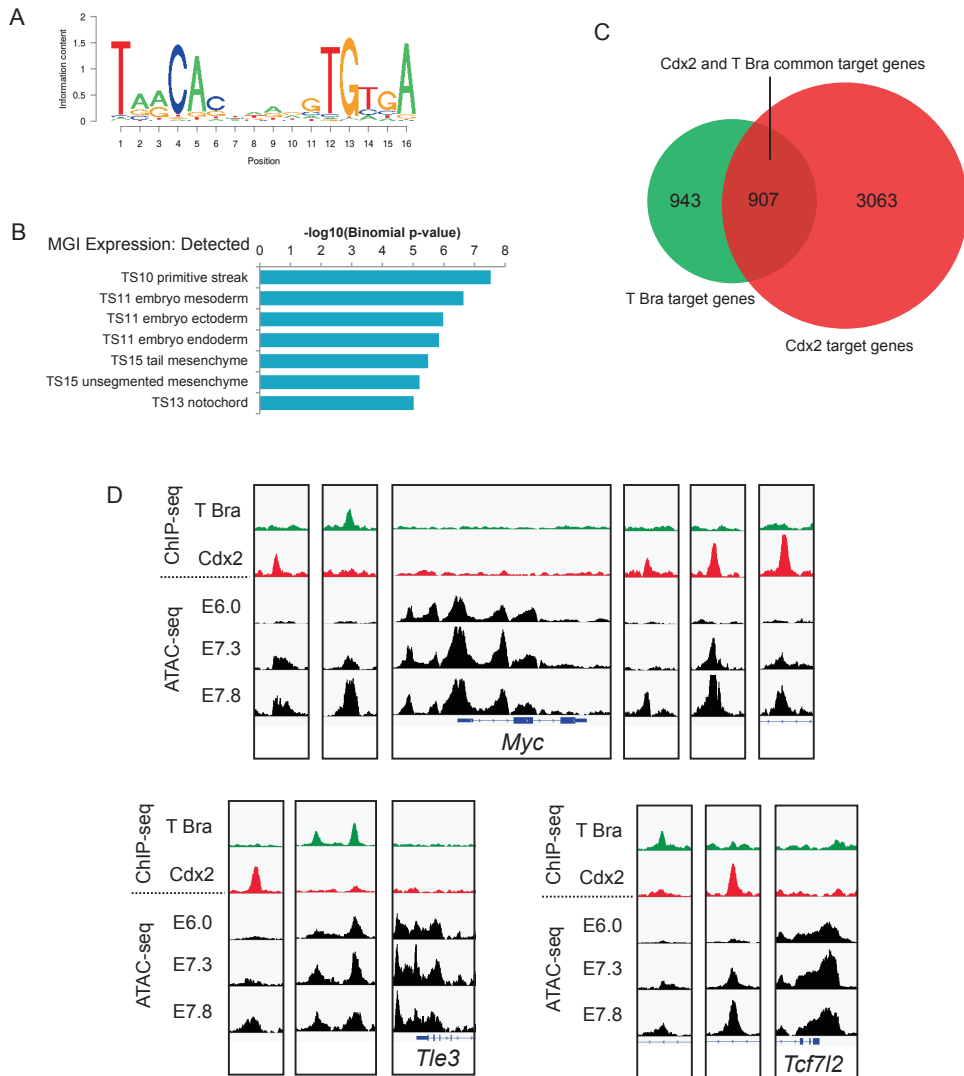


Figure 5. Cdx2 and T Bra bind common genes of the Wnt and Fgf signaling pathways. A) Sequence logo of the top enriched motif in T Bra-bound 200 bp summit regions identified in induced EpiSCs by SeqPos motif tool in Galaxy Cistrome. B) Top overrepresented ‘MGI Expression: Detected’ categories identified by GREAT analysis of T Bra ChIP-seq binding regions. The length of the bars corresponds to the binomial raw (uncorrected) p-value (x-axis). TS, Theiler stage. C) Overlap of Cdx2 and T Bra-bound genes identified by GREAT basal plus extension gene association rule. D) ATAC-seq profiles at *Myc*, *Tle3* and *Tcf712* loci in embryos at increasing developmental stages. See also Figure S5.

among these common targets are several genes belonging to the Wnt and Fgf pathways, including Wnt ligands (Figure S5D, Table S6). These were all associated by at least some criteria with activation of their target gene expression in embryos and in EpiSCs (Figure S5D). Upon examination of the list of these common targets of Cdx2 and T Bra together versus Cdx2 and T Bra alone (Table S6), it appears that Cdx2 binds many more loci in the Wnt and Fgf pathways than T Bra does. Notably T Bra and Cdx2 bind to each other's gene locus (Table S6) without affecting each other's transcription in single mutant embryos (Chawengsaksophak et al., 2004; Lolas et al., 2014; van de Ven et al., 2011). Nevertheless, according to our RNA-seq data and previous work, *T Bra* is significantly downregulated in Cdx triple mutant embryos (Table S3) (van Rooijen et al., 2012).

Some loci binding both Cdx2 and T Bra exhibit binding of these factors at non-overlapping sites, like *Myc*, *Tle3* and *Tcf712* (Figure 5D). For these genes, ATAC-seq experiments demonstrated that chromatin accessibility increases at both Cdx2 and T Bra peaks in embryos at stages when NMPs contribute to trunk axial extension compared to early embryos (Figure 5D). A subset of the genes that bind both Cdx2 and T Bra do so in regions that also exhibit enhancer properties (H3K27ac enrichment and/or open chromatin) (Figure 6A).

Loci that bind T Bra only were *Foxa2* (found to bind T Bra in human as well) (Faial et al., 2015), and *Mesp1*, *Mesp2*, *Ripply2* and *Tbx6*. The latter targets concern an independent additional function for T Bra in regulating mesoderm fate, which Cdx2 does not share (Figure S5E). This explains the difference in appearance of the posterior tissues when T Bra is inactive.

To functionally validate the involvement of commonly bound regions in genes of the Wnt and Fgf pathways by T Bra and Cdx2 *in vivo*, we performed *lacZ* reporter assays to test their activity. These assays confirmed that the T Bra and Cdx2 co-bound regions near *Rspo3*, *Fgf8* and *Fgf4-Fgf15* specifically activate gene transcription in the posterior region of E9.5 embryos (Figure 6B). Interestingly, Cdx2 binding sites identified by our Cdx2 ChIP-seq were more conserved than T Bra binding sites, and common binding sites were highly conserved (Figure S6A). The sequences corresponding to the peaks that were tested in *lacZ* assays were also found to be strongly conserved evolutionarily between human and mouse (Figure S6B). Validation of Cdx2 and T Bra enrichment at these latter loci was performed by ChIP-qPCR in embryonic tailbud versus anterior tissues at E9.5 (Figure 6C). Binding of both Cdx2 and T Bra to regions corresponding to the genes of the Wnt and Fgf pathways discovered in our ChIP-seq experiments was found to be exclusive for posterior embryonic tissues (Figure 6C). The data collectively demonstrate that Cdx2 and T Bra bind together to genes of the Wnt and Fgf pathway and that this binding causes target gene activation, a prerequisite for maintenance of the posterior progenitor niche in embryos.

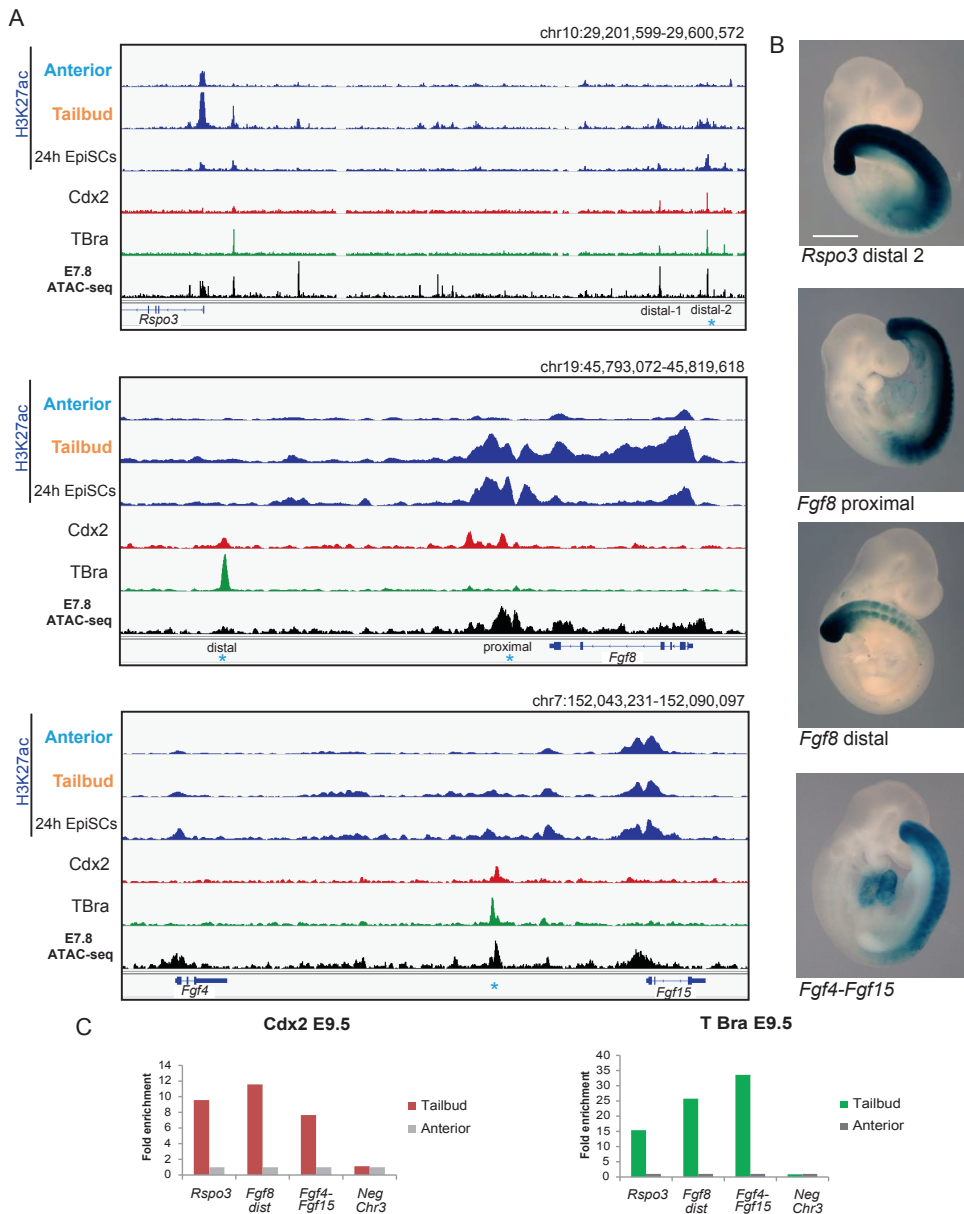


Figure 6. Cdx2 and T Bra co-bound regions activate genes of the Wnt and Fgf signaling pathway *in vivo*. A) ChIP-seq for H2K27ac, Cdx2 and T Bra and ATAC-seq for E7.8 embryos in the genomic regions containing *Rspo3* (top panels), *Fgf8* (middle panels) and *Fgf4-Fgf15* (bottom panels). Embryo tracks are in bold. Asterisks are regions further tested in *lacZ* transgenic assays. B) Activity of Cdx2 and T Bra co-bound regions coupled to *lacZ* reporter in E9.5 WT embryos. Top panel is *Rspo3* distal enhancer 2, middle panels are *Fgf8* proximal enhancer and *Fgf8* distal enhancer, and bottom panel is *Fgf4-Fgf15* enhancer. Anterior is on top. Scale bar is 1mm. C) Cdx2 and T Bra occupancy by ChIP-qPCR in posterior versus anterior E9.5 embryonic tissues. Enrichment of each region following immunoprecipitation with anti-Cdx2 and anti-T Bra antibody is calculated as fold enrichment over negative anterior embryonic tissue. Neg chr3 is a negative control region. See also Figure S6.

Discussion

Cdx, *T Bra* and axial extension from posterior progenitors

In the mouse embryo, trunk tissues are formed from populations of progenitors arising within the epiblast at the beginning of somitogenesis. Progenitors located between the node and the anterior end of the primitive streak, called the NSB, have been shown to constitute a population of stem cell-like NMPs (Cambray and Wilson, 2002, 2007) that contribute descendants to long axial distances. These NMPs were defined by co-expressing *T Bra* and *Sox2*. Epiblast cells abutting the NMP region posteriorly on both sides of the primitive streak have been called the caudo-lateral epiblast (CLE) (Wilson et al., 2009). The progenitors in the caudal CLE and NMP regions differ in their dependence on signaling pathways. Canonical Wnt was shown to regulate the size of the NMP population (Wymeersch et al., 2016) and is believed to maintain this progenitor pool, whereas the mesoderm progenitors in the posterior CLE are less dependent on this signaling. These data set the stage on which the action of *T Bra* and *Cdx2* play their role in ensuring trunk and posterior axial extension from the posterior progenitor populations by sustaining Wnt and Fgf signaling in the NMP niche. The co-expression of *Cdx* and *T Bra* in NMPs and their niche, the *Cdx2* and *T Bra* mutant phenotypes, their rescue by Wnt, and the binding of *Cdx* and *T Bra* at Wnt pathway loci, all suggest that these transcription factors regulate the NMPs and anterior CLE populations of progenitors. Besides NMP progenitor pool amplification and maintenance, Wnt signaling plays an additional role in driving the differentiation choice of NMPs. The *T Bra* and *Cdx2* loss of function and precocious expression of *Hox13* affect this differentiation choice of the NMPs as shown by the fact that in all these situations, ectopic neural tissues are formed posteriorly in embryos. The data collectively establish *Cdx2* expression as an essential factor in the signature of posterior NMP progenitors, together with *T Bra*. Interestingly, *T Bra/ntl* was reported to bind to genes specifying posterior fate in zebrafish embryos (Morley et al., 2009), one of which was *cdx4*. *T Bra* and *Cdx* therefore may belong to an evolutionary conserved mechanism driving NMP development. In addition, our data indicate that *Cdx2* is likely to act as a pioneer transcription factor on genes that maintain the niche of these NMP progenitors.

Both *Cdx2* and *T Bra* directly activate genes sustaining posterior axial elongation in NMPs and their niche

Cdx2 and *T Bra* proteins are essential for NMP maintenance in the embryonic posterior growth zone. The NMP niche is dependent on both Wnt and Fgf, as proven from the posterior truncation phenotype of mutants in the Wnt (Galceran et al., 1999; Takada et al., 1994) and Fgf pathways (Hoch and Soriano, 2006; Naiche et al., 2011). The functional activation of Wnt and Fgf pathway loci upon binding of *Cdx2* and *T Bra* indicates that these transcription factors are actively contributing to axial growth at the time embryos generate their trunk. Our CHIP-seq experiments identify some of the ligands and agonists of the Wnt and Fgf pathways that are directly targeted by both *Cdx2* and *T Bra*. As reported for adult intestinal stem cells

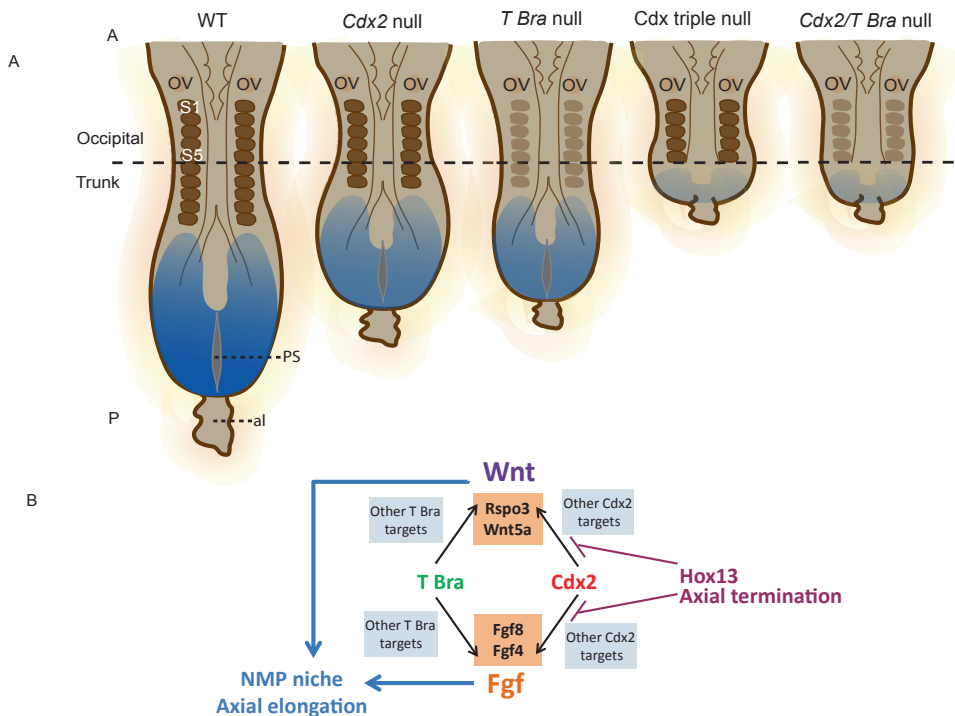


Figure 7. Schematic representation of the loss of Wnt and Fgf signaling in *Cdx* and *T Bra* mutant embryos and model for maintenance of the niche. A) Schematic dorsal view of E8.5 (left to right) WT, *Cdx2* null, *T Bra* null, *Cdx* triple null and *Cdx2* null/*T Bra* null mutant embryos. Wnt and Fgf signaling gradient is in blue. Somites are formed irregularly in *T Bra* mutants (light brown). Al, allantois; OV, otic vesicle; PS, primitive streak; S, somite. A, anterior and P, posterior. B) Model of *Cdx2*, *T Bra*, Wnt and Fgf gene regulatory network to maintain the niche for axial extension. Hox13 binds to *Cdx2* targets to antagonize *Cdx2* action and contributes to axial termination.

(Clevers et al., 2014), we conclude that *Cdx2* and *T Bra*-expressing NMPs stimulate and maintain their own niche.

Collaboration of *Cdx* and *T Bra* in embryonic trunk morphogenesis

The *Cdx* genes are required for the generation of the complete post-occipital part of the axis. This is in line with *Cdx* genes being paraHox genes (sharing ancestry with the Hox genes), being exclusively involved in the generation and patterning of the trunk and posterior tissues. Despite the fact that *T Bra* is expressed much earlier than when NMPs start contributing to trunk elongation (early somite stage (Wymeersch et al., 2016)), *T Bra* null mutants generate a normal anterior part of their axis including occipital tissues (the first 7 somites). It has been proposed that other T-box genes that bind the same motif as *T Bra*, possibly account for

the axial growth during the generation of the anterior set of somites in *T Bra* null mutants (Gentsch et al., 2013). Whatever it may be, *Cdx2* loss of function reduces the amount of posterior axial tissues made in *T Bra* mutants to the remaining occipital section of the axis in the compound mutants. The fact that the double mutants lacking both *Cdx2* and *T Bra* exhibit a severe posterior truncation at the same axial level as the *Cdx* null mutants (Figure S4C) suggests that the action of *Cdx2* on posterior tissue generation is similarly supplemented by either *Cdx1* and *Cdx4* as shown before (van Rooijen et al., 2012), or by *T Bra*. This results from the fact that two partially redundant systems, *T Bra* and *Cdx* converge in their output on the growth signaling activity in the axial progenitor niche. This is illustrated in Figure 7A.

Cdx, *T Bra* and the Hox genes instruct the trunk to tail transition and the termination of axial growth

T Bra and *Cdx2* are indispensable driving forces for trunk axis extension. Previous work showed that genes in the middle region of the *Hox* clusters could assist *Cdx* in supporting trunk axial growth (Young et al., 2009). We found that *Cdx2* binds to many sites in the anterior and middle part of the *Hox* clusters in EpiSCs and in embryos, and upregulates these genes in both these systems. This upregulation added to the autoregulation of *Cdx2* and to its cross-regulation by *T Bra*, insures a robust output of *Cdx* and *Hox* stimulators on trunk axial growth. The situation changes after the trunk to tail transition. *Cdx* genes are downregulated by E12.5 and *T Bra* by E14.5. In addition, the posterior *Hox* genes such as *Hox13*, which are turned on later than the anterior and central genes of the clusters, are expressed at high levels in the growth zone after the trunk to tail transition. *Hox13* gene products have been shown to repress axial growth when expressed precociously, antagonizing the action of anterior and central *Hox* and *Cdx* genes (Young et al., 2009). A repressive function of *Hox13* on more anterior *Hox* genes was recently demonstrated mechanistically for *HoxA* genes in the limb (Beccari et al., 2016). Due to this repressive function, *Hox13* proteins would interrupt the transcription of any anterior and central *Hox* gene that would still be expressed in the growth zone after the trunk to tail transition. This negative regulation of *Hox13* on more anterior *Hox* genes would occur spatially in the region where NMPs are located. In addition, our data show binding of *Hox13* and *Cdx2* on the same genes of the growth signaling pathways. We therefore propose that the *Cdx2/Hox* trunk-stimulating loop is weakened at the trunk to tail transition by the dual action of *Hox13* proteins that on one end repress more anterior *Hox* genes and on the other end antagonize *Cdx2* binding. In addition to the downregulation of *T Bra* and *Cdx* (Young et al., 2009), this reduces Wnt and Fgf signaling in the growth zone, exhausting the niche of the axial progenitors NMPs, and foreshadowing the termination of axial elongation (Figure 7A).

This study establishes *Cdx2* as a key transcriptional component that collaborates with *T Bra* in the maintenance of the NMPs and their niche during trunk axial elongation in embryos (Figure 7B). In addition to being downregulated after the trunk/tail transition, *Cdx2* is counteracted by *Hox13* proteins ensuring axis extension termination.

Material and methods

Mouse strains

All experiments using mice were performed in accordance with the institutional and national guidelines and regulations, under the control of the Dutch Committee for Animals in Experiments. All mice were in the C57Bl6j/CBA mixed background. The *Cdx2* conditional allele was described previously (Stringer et al., 2012) and the *T Brachyury*^{2l} deletion allele was described previously (Herrmann et al., 1990). Details of all transgenic mice are described in Supplemental Experimental Procedures. ISH of transcripts in mouse embryos was performed as described previously (Young et al., 2009).

EpiSCs derivation and culture

WT and *Cdx* triple null pre-streak embryos (E6.0) were isolated in M2 medium; extraembryonic tissue and surrounding primitive endoderm was removed as described previously (Neijts et al., 2016). Detailed cell culture and induction procedures are described in Supplemental Experimental Procedures.

RNA-seq

Mouse embryos were isolated at E8.0 as described previously (Young et al., 2009), 8 dissected tailbuds were pooled for each replicate (3 to 5 somites stage). For mRNA sequencing 10 ng of total RNA was used as starting material and was processed using the CEL-Seq protocol (Hashimshony et al., 2012) described in Supplemental Experimental Procedures.

ChIP and ChIP-seq

ChIP-seq was performed on EpiSCs according to a standard ChIP protocol. Detailed ChIP-seq analysis is described in Supplemental Experimental Procedures. ChIP on 50 E9.5 tailbud and anterior mouse embryonic tissue was performed as described previously (Amin and Bobola, 2014). For differential posterior versus anterior E9.5 embryonic tissues, posterior tissues were the posterior parts of the embryos dissected at the level of the last formed somite, and anterior tissues were taken between the branchial arches and forelimb bud. Antibodies are described in Supplemental Experimental Procedures.

Assay for Transposase-Accessible Chromatin (ATAC)-seq

ATAC-seq was performed as described previously (Neijts et al., 2016).

Statistical analysis

For ChIP-seq experiments, statistically significant enriched regions for *Cdx2*, *T Bra* and H3K27ac were identified using MACS (Zhang et al., 2008) with a p-value threshold = 10⁻⁵. For RNA-seq experiments statistically significant differentially expressed genes were identified using DESeq2 (Love et al., 2014). An FDR < 0.05 was used. For heterogeneous embryonic

material, genes with a p-value < 0.05 were identified.

Accession numbers

All data are deposited on GEO under accession number GSE84899 and GSE81203.

Author contribution

J.D, S.A., R.N. and S.S. conceived the study; S.A., R.N., S.S. and C.v R. performed the experiments. S.A. performed the ChIP experiments, RNA-seq experiments and NGS data analysis. R.N. generated and characterized EpiSCs and performed the ATAC-seq experiments, S.S. generated the Cdx2/T Bra double mutant embryos. C.v.R. generated the Cdx triple null mice and performed ISH experiments. L.K. processed the samples for RNA-seq, according to the technology developed in the lab of A. v O. S.T. performed the bioinformatics analyses of ChIP-seq, RNA-seq and ATAC-seq data; M.P.C. provided expert advice on some of the ChIP-seq experiments. S.A. and J.D. wrote the manuscript.

Acknowledgments

We thank Jeroen Korving for micro-injection of the *lacZ* reporter constructs. We thank De-nen Wellik for providing us with the *Hoxb13-FLAG* expression construct, and Joris Slingerland for characterizing the homozygous animals of our *Cdx2P-Hoxb13-FLAG* mouse colonies. We thank Guilherme Costa and Wim de Graaff for some of the ISH, and Fabrizio Giuliani for help with immuno-staining in one of the experiments. We thank Maartje Vermunt for help with bioinformatics analysis. This work was supported by a grant from the Netherlands Institute for Regenerative Medicine (NIRM, grant FES0908). The authors declare no conflict of interest.

References

- Amin, S., and Bobola, N. (2014). Chromatin immunoprecipitation and chromatin immunoprecipitation with massively parallel sequencing on mouse embryonic tissue. *Methods Mol Biol* 1196, 231-239.
- Beccari, L., Yakushiji-Kaminatsui, N., Woltering, J.M., Necsulea, A., Lonfat, N., Rodriguez-Carballo, E., Mascrez, B., Yamamoto, S., Kuroiwa, A., and Duboule, D. (2016). A role for HOX13 proteins in the regulatory switch between TADs at the HoxD locus. *Genes & development* 30, 1172-1186.
- Benahmed, F., Gross, I., Gaunt, S.J., Beck, F., Jehan, F., Domon-Dell, C., Martin, E., Kedinger, M., Freund, J.N., and Duluc, I. (2008). Multiple regulatory regions control the complex expression pattern of the mouse Cdx2 homeobox gene. *Gastroenterology* 135, 1238-1247, 1247 e1231-1233.
- Bialecka, M., Wilson, V., and Deschamps, J. (2010). Cdx mutant axial progenitor cells are rescued by grafting to a wild type environment. *Developmental biology* 347, 228-234.
- Cambray, N., and Wilson, V. (2002). Axial progenitors with extensive potency are localised to the mouse chordoneural hinge. *Development* 129, 4855-4866.
- Cambray, N., and Wilson, V. (2007). Two distinct sources for a population of maturing axial progenitors. *Development* 134, 2829-2840.
- Chawengsaksophak, K., de Graaff, W., Rossant, J., Deschamps, J., and Beck, F. (2004). Cdx2 is essential for axial elongation in mouse development. *Proceedings of the National Academy of Sciences of the United States of America* 101, 7641-7645.
- Chawengsaksophak, K., James, R., Hammond, V.E., Kontgen, F., and Beck, F. (1997). Homeosis and intestinal tumours in Cdx2 mutant mice. *Nature* 386, 84-87.
- Clevers, H., Loh, K.M., and Nusse, R. (2014). Stem cell signaling. An integral program for tissue renewal and regeneration: Wnt signaling and stem cell control. *Science* 346, 1248012.
- Denans, N., Iimura, T., and Pourquie, O. (2015). Hox genes control vertebrate body elongation by collinear Wnt repression. *eLife* 4.
- Faial, T., Bernardo, A.S., Mendjan, S., Diamanti, E., Ortmann, D., Gentsch, G.E., Mascetti, V.L., Trotter, M.W., Smith, J.C., and Pedersen, R.A. (2015). Brachyury and SMAD signalling collaboratively orchestrate distinct mesoderm and endoderm gene regulatory networks in differentiating human embryonic stem cells. *Development* 142, 2121-2135.
- Galceran, J., Farinas, I., Depew, M.J., Clevers, H., and Grosschedl, R. (1999). Wnt3a^{-/-}-like phenotype and limb deficiency in Lef1^(-/-)Tcf1^(-/-) mice. *Genes & development* 13, 709-717.
- Gentsch, G.E., Owens, N.D., Martin, S.R., Piccinelli, P., Faial, T., Trotter, M.W., Gilchrist, M.J., and Smith, J.C. (2013). In vivo T-box transcription factor profiling reveals joint regulation of embryonic neuromesodermal bipotency. *Cell reports* 4, 1185-1196.
- Gouti, M., Tsakiridis, A., Wymeersch, F.J., Huang, Y., Kleinjung, J., Wilson, V., and Briscoe, J. (2014). In vitro generation of neuromesodermal progenitors reveals distinct roles for wnt signalling in the specification of spinal cord and paraxial mesoderm identity. *PLoS biology* 12, e1001937.
- Hashimshony, T., Wagner, F., Sher, N., and Yanai, I. (2012). CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell reports* 2, 666-673.
- Herrmann, B.G., Labeit, S., Poustka, A., King, T.R., and Lehrach, H. (1990). Cloning of the T gene required in mesoderm formation in the mouse. *Nature* 343, 617-622.
- Hoch, R.V., and Soriano, P. (2006). Context-specific requirements for Fgfr1 signaling through Frs2 and Frs3 during mouse development. *Development* 133, 663-673.
- Kispert, A., and Herrmann, B.G. (1993). The Brachyury gene encodes a novel DNA binding protein. *The EMBO journal* 12, 3211-3220.
- Lolas, M., Valenzuela, P.D., Tjian, R., and Liu, Z. (2014). Charting Brachyury-mediated developmental pathways during early mouse embryogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 111, 4478-4483.

- Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* 15, 550.
- Martin, B.L. (2016). Factors that coordinate mesoderm specification from neuromesodermal progenitors with segmentation during vertebrate axial extension. *Seminars in cell & developmental biology* 49, 59-67.
- Martin, B.L., and Kimelman, D. (2008). Regulation of canonical Wnt signaling by Brachyury is essential for posterior mesoderm formation. *Developmental cell* 15, 121-133.
- Martin, B.L., and Kimelman, D. (2009). Wnt signaling and the evolution of embryonic posterior development. *Current biology* : CB 19, R215-219.
- Martin, B.L., and Kimelman, D. (2010). Brachyury establishes the embryonic mesodermal progenitor niche. *Genes & development* 24, 2778-2783.
- McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nature biotechnology* 28, 495-501.
- Morley, R.H., Lachani, K., Keefe, D., Gilchrist, M.J., Flicek, P., Smith, J.C., and Wardle, F.C. (2009). A gene regulatory network directed by zebrafish No tail accounts for its roles in mesoderm formation. *Proceedings of the National Academy of Sciences of the United States of America* 106, 3829-3834.
- Naiche, L.A., Holder, N., and Lewandoski, M. (2011). FGF4 and FGF8 comprise the wavefront activity that controls somitogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 108, 4018-4023.
- Neijts, R., Amin, S., van Rooijen, C., Tan, S., Creyghton, M.P., de Laat, W., and Deschamps, J. (2016). Polarized regulatory landscape and Wnt responsiveness underlie Hox activation in embryos. *Genes & development* 30, 1937-1942.
- Neijts, R., Simmini, S., Giuliani, F., van Rooijen, C., and Deschamps, J. (2014). Region-specific regulation of posterior axial elongation during vertebrate embryogenesis. *Developmental dynamics* : an official publication of the American Association of Anatomists 243, 88-98.
- Olivera-Martinez, I., Harada, H., Halley, P.A., and Storey, K.G. (2012). Loss of FGF-dependent mesoderm identity and rise of endogenous retinoid signalling determine cessation of body axis elongation. *PLoS biology* 10, e1001415.
- Rashbass, P., Wilson, V., Rosen, B., and Beddington, R.S. (1994). Alterations in gene expression during mesoderm formation and axial patterning in Brachyury (T) embryos. *The International journal of developmental biology* 38, 35-44.
- Rivera-Perez, J.A., and Hadjantonakis, A.K. (2015). The Dynamics of Morphogenesis in the Early Mouse Embryo. *Cold Spring Harbor perspectives in biology* 7.
- Rivera-Perez, J.A., and Magnuson, T. (2005). Primitive streak formation in mice is preceded by localized activation of Brachyury and Wnt3. *Developmental biology* 288, 363-371.
- Stringer, E.J., Duluc, I., Saandi, T., Davidson, I., Bialecka, M., Sato, T., Barker, N., Clevers, H., Pritchard, C.A., Winton, D.J., et al. (2012). Cdx2 determines the fate of postnatal intestinal endoderm. *Development* 139, 465-474.
- Takada, S., Stark, K.L., Shea, M.J., Vassileva, G., McMahon, J.A., and McMahon, A.P. (1994). Wnt-3a regulates somite and tailbud formation in the mouse embryo. *Genes & development* 8, 174-189.
- Takemoto, T., Uchikawa, M., Yoshida, M., Bell, D.M., Lovell-Badge, R., Papaioannou, V.E., and Kondoh, H. (2011). Tbx6-dependent Sox2 regulation determines neural or mesodermal fate in axial stem cells. *Nature* 470, 394-398.
- Tsakiridis, A., Huang, Y., Blin, G., Skylaki, S., Wymeersch, F., Osorno, R., Economou, C., Karagianni, E., Zhao, S., Lowell, S., et al. (2014). Distinct Wnt-driven primitive streak-like populations reflect in vivo lineage precursors. *Development* 141, 1209-1221.
- Tsakiridis, A., and Wilson, V. (2015). Assessing the bipotency of in vitro-derived neuromesodermal progenitors. *F1000Research* 4, 100.
- Turner, D.A., Hayward, P.C., Baillie-Johnson, P., Rue, P., Broome, R., Faunes, F., and Martinez Arias, A. (2014). Wnt/beta-catenin and FGF signalling direct the specification and maintenance of a neuromesodermal axial progenitor in ensembles of mouse embryonic stem cells. *Development* 141, 4243-4253.

- Tzouanacou, E., Wegener, A., Wymeersch, F.J., Wilson, V., and Nicolas, J.F. (2009). Redefining the progression of lineage segregations during mammalian embryogenesis by clonal analysis. *Developmental cell* 17, 365-376.
- van de Ven, C., Bialecka, M., Neijts, R., Young, T., Rowland, J.E., Stringer, E.J., Van Rooijen, C., Meijlink, F., Novoa, A., Freund, J.N., et al. (2011). Concerted involvement of Cdx/Hox genes and Wnt signaling in morphogenesis of the caudal neural tube and cloacal derivatives from the posterior growth zone. *Development* 138, 3451-3462.
- van den Akker, E., Forlani, S., Chawengsaksophak, K., de Graaff, W., Beck, F., Meyer, B.I., and Deschamps, J. (2002). Cdx1 and Cdx2 have overlapping functions in anteroposterior patterning and posterior axis elongation. *Development* 129, 2181-2193.
- van Rooijen, C., Simmini, S., Bialecka, M., Neijts, R., van de Ven, C., Beck, F., and Deschamps, J. (2012). Evolutionarily conserved requirement of Cdx for post-occipital tissue emergence. *Development* 139, 2576-2583.
- Visel, A., Minovitsky, S., Dubchak, I., and Pennacchio, L.A. (2007). VISTA Enhancer Browser--a database of tissue-specific human enhancers. *Nucleic acids research* 35, D88-92.
- Wilkinson, D.G., Bhatt, S., and Herrmann, B.G. (1990). Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature* 343, 657-659.
- Wilson, V., and Beddington, R. (1997). Expression of T protein in the primitive streak is necessary and sufficient for posterior mesoderm movement and somite differentiation. *Developmental biology* 192, 45-58.
- Wilson, V., Olivera-Martinez, I., and Storey, K.G. (2009). Stem cells, signals and vertebrate body axis extension. *Development* 136, 1591-1604.
- Wymeersch, F.J., Huang, Y., Blin, G., Cambray, N., Wilkie, R., Wong, F.C., and Wilson, V. (2016). Position-dependent plasticity of distinct progenitor types in the primitive streak. *eLife* 5.
- Yamaguchi, T.P., Bradley, A., McMahon, A.P., and Jones, S. (1999). A Wnt5a pathway underlies outgrowth of multiple structures in the vertebrate embryo. *Development* 126, 1211-1223.
- Yoshikawa, Y., Fujimori, T., McMahon, A.P., and Takada, S. (1997). Evidence that absence of Wnt-3a signaling promotes neuralization instead of paraxial mesoderm development in the mouse. *Developmental biology* 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., et al. (2009). Cdx and Hox genes differentially regulate posterior axial growth in mammalian embryos. *Developmental cell* 17, 516-526.
- Zaret, K.S., and Mango, S.E. (2016). Pioneer transcription factors, chromatin dynamics, and cell fate control. *Current opinion in genetics & development* 37, 76-81.
- Zhang, Y., Liu, T., Meyer, C.A., Eeckhoutte, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome biology* 9, R137.

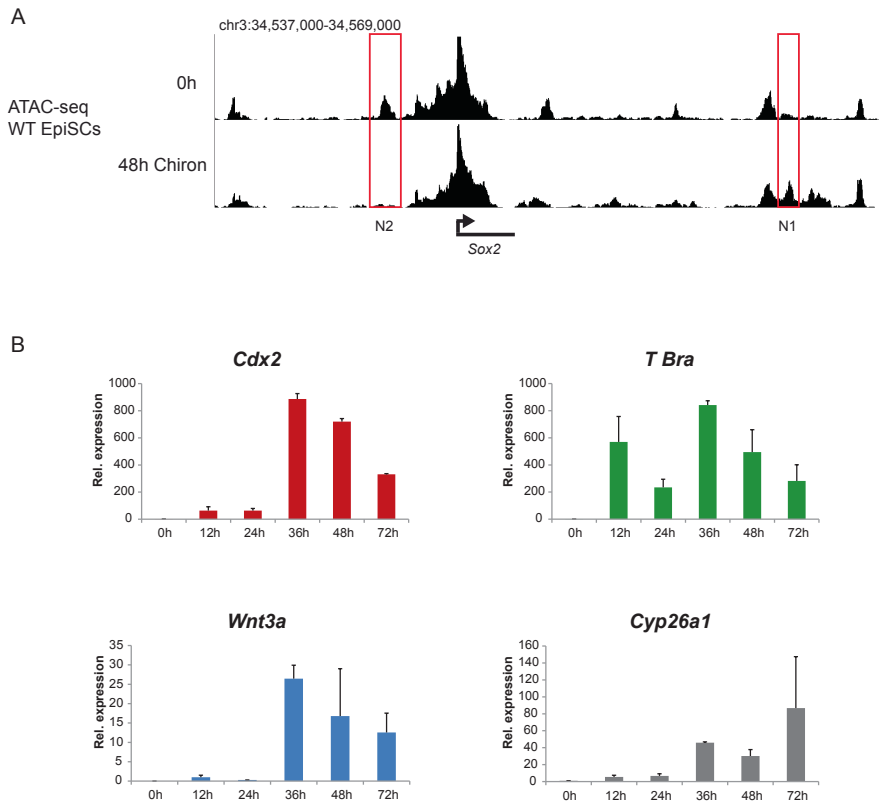


Figure S1. Characterization of induced WT EpiSCs and induction of posterior markers upon administration of *Wnt3a*. Related to Figure 1. A) ATAC-seq profile corresponding to the genomic region containing *Sox2* *N1* enhancer in uninduced and induced WT EpiSCs. B) Induction time course of *Cdx2* expression and markers for posterior axial extension upon *Wnt3a* stimulation. Values represent expression levels relative to 0 hours and are normalized to *Ppia*. Error bars represent SD of at least two biological replicates. h, hours.

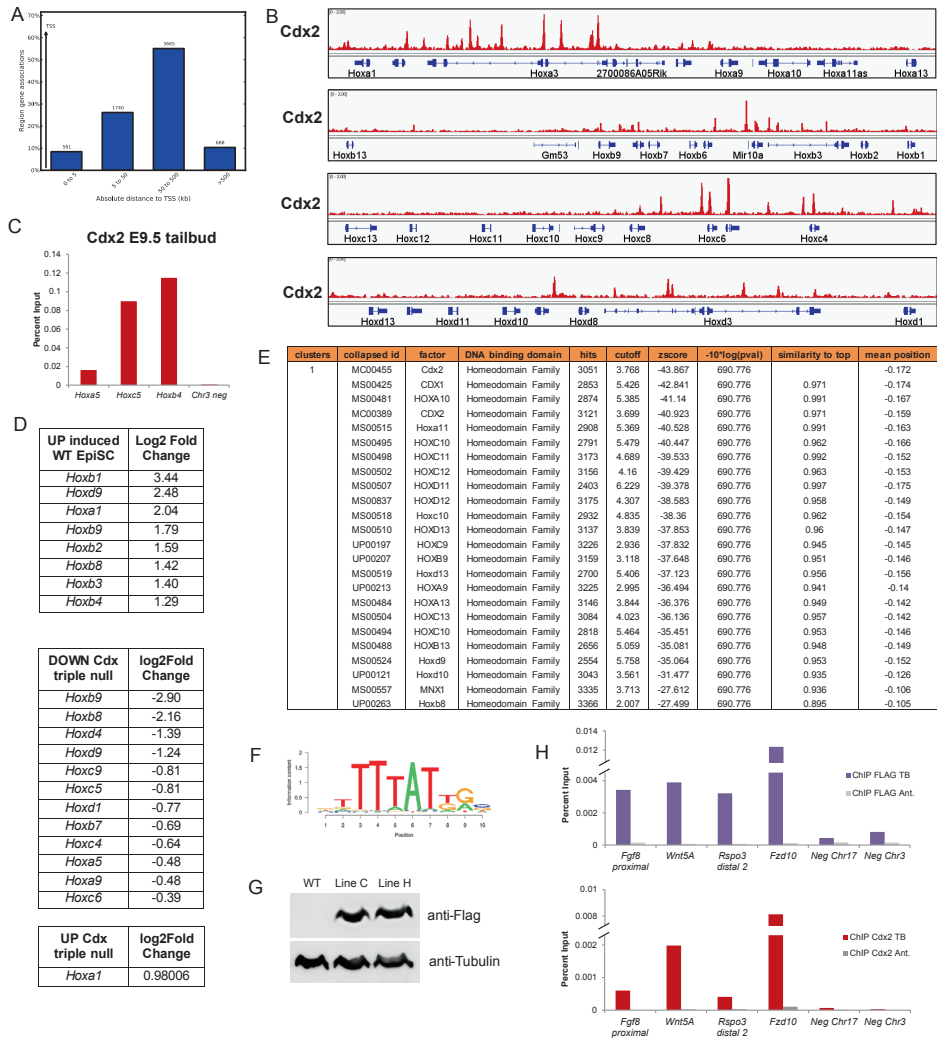


Figure S2. Cdx2 regulation of middle Hox genes and commonly occupied regions of posterior Hox genes and Cdx2. Related to Figure 2. A) Distribution of Cdx2 binding regions relative to the absolute distance to a transcription start site (TSS) identified by GREAT. B) ChIP-seq tracks corresponding to Cdx2 binding at the Hox clusters. C) Cdx2 occupancy at *Hoxa5*, *Hoxc5* and *Hoxb4* by ChIP-qPCR in E9.5 embryonic tailbud tissue. Enrichment of each region following immunoprecipitation with anti-Cdx2 antibody is calculated as percent input. Chr3 neg is a negative control region. D) List of Hox genes affected in induced WT EpiSCs (top) and Cdx Triple null vs WT embryos (middle and bottom) by RNA-seq. E) List of top enriched mouse and human DNA motifs identified by SeqPos motif tool in Galaxy Cistrome. F) HOXB13 motif identified by SeqPos. G) Western blot analysis of WT and Flag-Hoxb13 dissected E9.5 tailbud extracts using anti-Flag and anti-tubulin antibodies. H) Cdx2 and Hoxb13 occupancy by ChIP-qPCR in embryonic tailbud (TB) and anterior tissues (Ant). Enrichment of each region following immunoprecipitation with anti-Cdx2 and anti-Flag antibody is calculated as percent input. Neg chr3 and Neg chr17 are negative control regions.

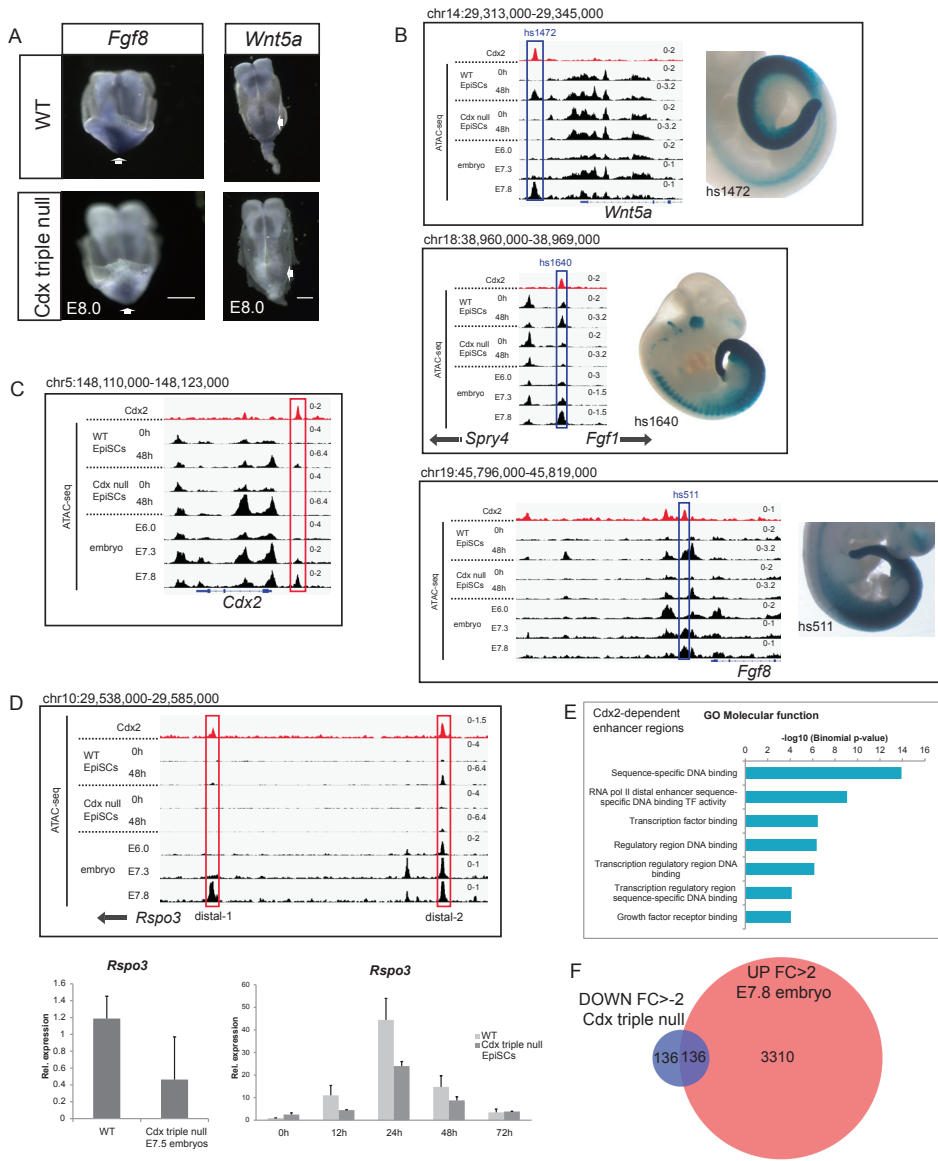


Figure S3. Validation of *Cdx2* targets. Related to Figure 3. A) Whole-mount ISH on E8.0 WT and Cdx triple null embryos using *Fgf8* and *Wnt5a* probes. Scale bar is 100 μ m. B) ATAC-seq profiles of VISTA validated enhancers (blue boxes) overlapping *Cdx2* binding regions corresponding to the genomic regions containing *Wnt5a*, *Spry4-Fgf1* and *Fgf8* in uninduced and induced WT and Cdx triple null EpiSCs, and in early embryos. C) ATAC-seq profile of the *Cdx2* promoter region in uninduced and induced WT and Cdx triple null EpiSCs, and in early embryos. *Cdx2* binding region is highlighted in the red box. D) ATAC-seq profile of *Rspo3* distal *Cdx2* binding regions (red boxes) in uninduced and induced WT and Cdx triple null EpiSCs, and in early embryos. *Rspo3* expression levels in WT and Cdx triple null early embryos. Values represent expression relative to internal control. Induction time course of *Rspo3* expression in WT and Cdx triple null EpiSCs upon Chiron stimulation for up to 72 hours. Values represent expression relative to WT EpiSCs at 0 hours and normalized to *Ppia*. Error bars represent SD of at least... (continue on next page)

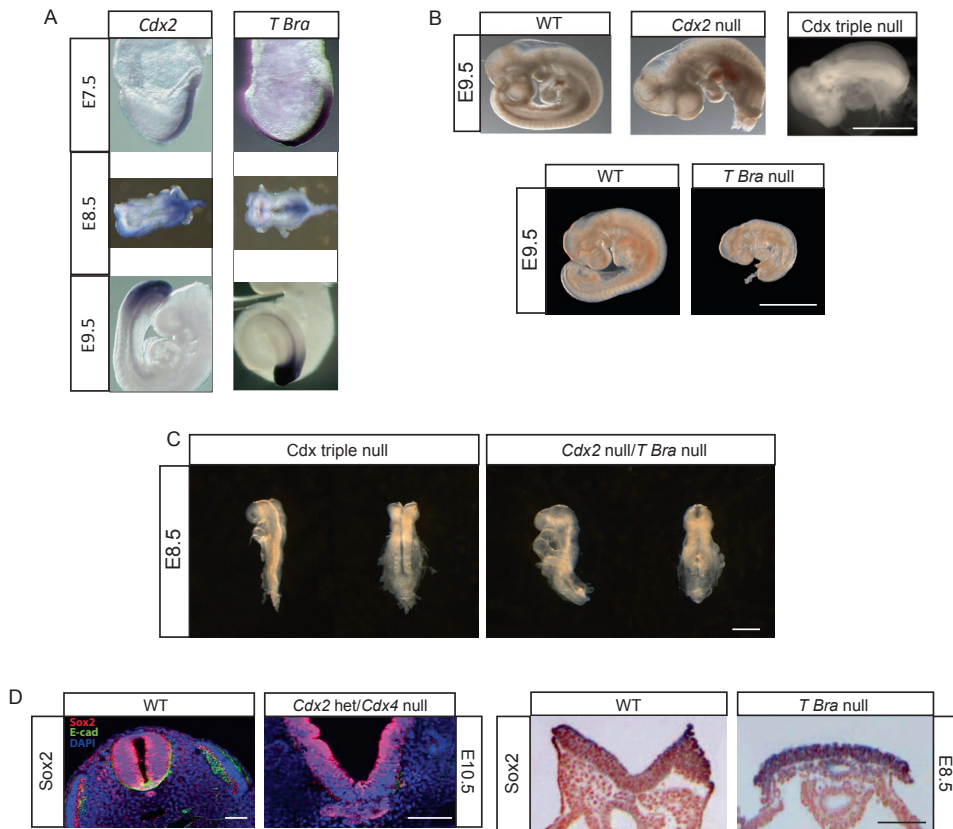


Figure S4. Expression of *Cdx2* and *T Bra* during axial elongation and ectopic neural structures at posterior levels in *T Bra* null embryos and *Cdx* compound mutants. Related to Figure 4. A) Expression of *Cdx2* and *T Bra* by whole-mount ISH in WT embryos at E7.5, E8.5 and E9.5. B) Top, allelic series of *Cdx* compound mutant embryos at E9.5. The severity of the posterior truncation progressively increases with the loss of *Cdx* alleles. Anterior is to the left. Bottom, *T Bra* null mutant compared with WT embryo at E9.5. *T Bra* null mutant embryo shows a severe posterior truncation. Scale bars are 2 mm. C) Phenotype comparison of *Cdx* triple null and *Cdx2* null/*T Bra* null in E8.5 mutant embryos. Scale bar is 500 μ m. D) (Left) *Sox2* expressing neural structures in transversal sections of the tailbud region of E10.5 *Cdx2* het/*Cdx4* null mutants compared with WT. Ectopic neural structures are formed ventral to the neural tube at posterior levels of the axis, where the neural tube is still open in mutants. Scale bar is 50 μ m. (Right) *Sox2* expression in transversal sections of the posterior region of E8.5 *T Bra* null mutants compared with WT. *T Bra* null mutants exhibit clear posterior ectopic neural structures. Scale bar is 100 μ m.

(continue from previous page) ...two biological replicates. h, hours. E) Top overrepresented molecular function categories identified by GREAT analysis of *Cdx2* ChIP-seq binding regions that overlap with regions inaccessible in induced *Cdx* triple null EpiSCs (FC>-1.5). The length of the bars corresponds to the binomial raw (uncorrected) P-value (x-axis). F) Overlap of top regions inaccessible in induced *Cdx* triple null EpiSCs (FC>-2) and accessible in WT E7.8 embryos (FC>2).

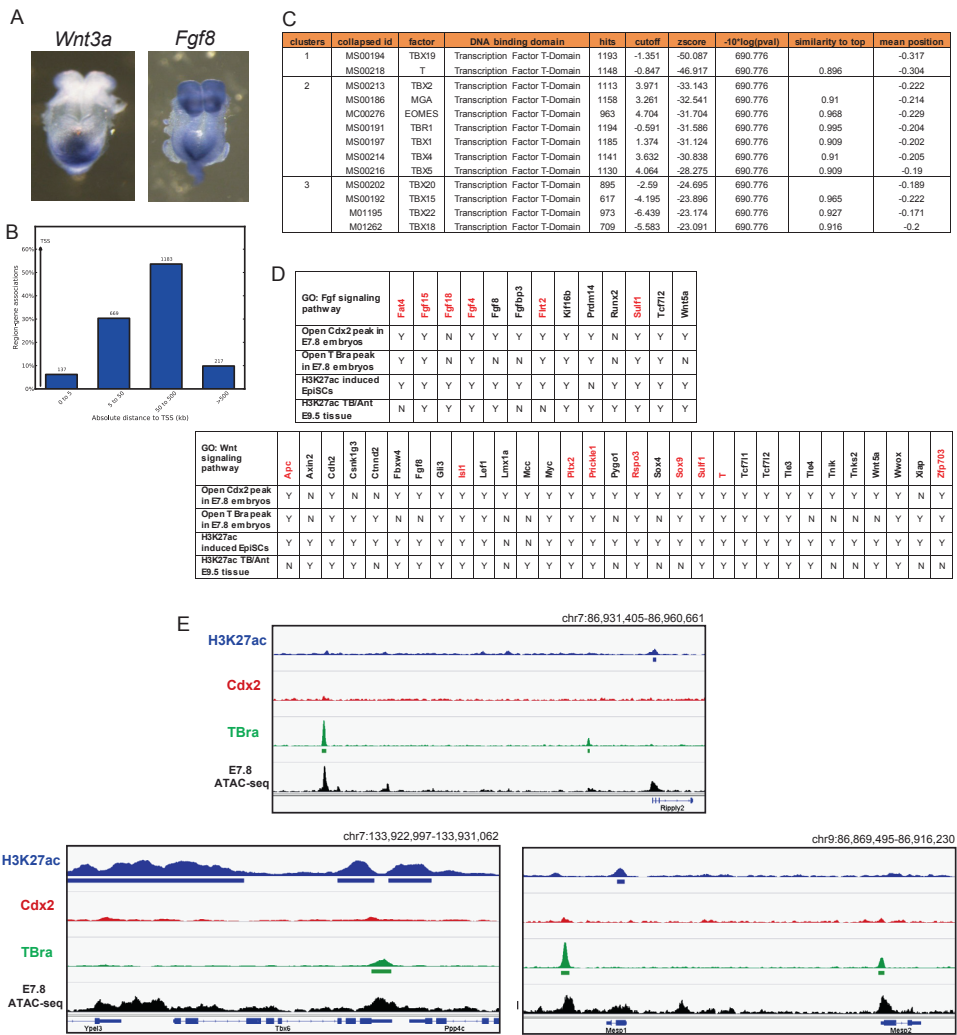


Figure S5. T Bra occupancy and functional validation. Related to Figure 5. A) Whole-mount ISH on E8.0 WT embryos using *Wnt3a* and *Fgf8* probes. B) Distribution of T Bra binding regions relative to the absolute distance to a transcription start site (TSS) identified by GREAT in the ChIP-seq experiment on induced EpiSCs. C) List of top motifs identified by SeqPos motif tool in Galaxy Cistrome. D) List of Cdx2 and T Bra common targets in the Wnt and Fgf signaling pathways. ATAC-seq identifies 'open' areas in E7.8 embryo chromatin; H3K27ac positive regions surrounding the ChIP-seq peak or gene to indicate active regions in induced EpiSCs and in E9.5 embryonic posterior tailbud versus anterior trunk tissue (TB/Ant). Red highlighted genes are genes with overlap in Cdx2 and T Bra binding in 3 kb region around the summit of peaks. TB, tailbud and Ant, anterior. Y, yes and N, no. E) ChIP-seq tracks for H3K27ac, Cdx2 and T Bra and ATAC-seq track for E7.8 embryo corresponding to the genomic regions containing *Mesp1-Mesp2*, *Tbx6* and *Ripply2*. Solid bars under each track represent the MACS peak calling identified regions.

Supplemental Tables

All Supplemental Tables are available on www.cell.com/cell-reports/ and are available from the author upon request.

Table S1. 24 h vs 0 h WT EpiSC RNA-seq.

Related to Figure 1. Genes upregulated in induced WT EpiSCs following 24 hours induction with Chiron and Fgf8.

Table S2. Cdx2 pooled ChIP-seq peaks WT EpiSCs 24 h induction.

Related to Figure 2. Genomic coordinates of Cdx2 binding regions in induced WT EpiSCs following 24 hours induction with Chiron and Fgf8.

Table S3. Embryo RNA-seq.

Related to Figure 2. Up and downregulated genes in Cdx2 triple vs WT embryos, *Cdx1-4* null vs WT embryos and Cdx triple null vs *Cdx1-4* null embryos.

Table S4. Cdx2-dependent enhancers.

Related to Figure 3. Regions associated to genes inaccessible in induced Cdx triple null EpiSCs vs induced WT EpiSCs and accessible in E7.8 embryos, bound by Cdx2.

Table S5. T Bra pooled ChIP-seq peaks WT EpiSCs 24 h induction.

Related to Figure 5. Genomic coordinates of T Bra binding regions in induced WT EpiSCs following 24 hours induction with Chiron and Fgf8.

Table S6. T Bra ChIP-seq and Cdx2 ChIP-seq common targets.

Related to Figure 5. Genes associated to ChIP-seq identified regions for T Bra and Cdx2 and their common target genes.

Supplemental Experimental Procedures

Mouse strains

Generation of the strain carrying *Cdx2* conditional allele was described previously (Stringer et al., 2012). The *Rosa25CreERT2* mice were obtained from Austin Smith, Wellcome Trust Centre for Stem Cell Research, University of Cambridge, UK. *Cdx2^{fl/fl}/RosaCreERT2* mice were generated by interbreeding *Cdx2^{fl/fl}* and *RosaCreERT2* mice. The generation of the strain carrying *T Brachyury^{2j}* deletion (deletion of 80-110 kb on chromosome 17) was described by Herrmann et al., 1990 (Herrmann et al., 1990). The generation of *Cdx* triple null strain was described previously (van Rooijen et al., 2012). *Cdx2* null/*T Bra* null strain was generated by interbreeding *Cdx2^{fl/fl}/RosaCreERT2/T Bra^{+/-}* mice. Pregnant females underwent intraperitoneal tamoxifen injection at E5.0 in order to activate Cre recombinase and to induce *Cdx2* null. *Cdx2P-Hoxb13* mice were generated as described previously (Young et al.).

A short or no tail phenotype was used to identify heterozygous mutant mice carrying the *T Bra^{2j}* deletion. Genotyping of *T Bra* null embryos was performed by using qPCR analysis to calculate the difference between *T Bra* and *Gapdh* and β -*actin* measurements. Primers used were: *T Bra* Fwd 5'-AGTCTGCAAAGCCCTGTGAT-3' and Rev 5'-ATCGGAGAACCAGAAGACGA-3'. *Gapdh* Fwd 5'-AGCTGATGGCTGCAGGTTCTC-3' and Rev 5'-GCCACTGCAAATGGCAGCCCT-3'. β -*actin* Fwd 5'-AGGCCCCAGCAACACGTCATT-3' and Rev 5'-GGCCCTTCATTGTGGCTGCCT-3'.

EpiSCs culture and induction

Cdx triple null EpiSCs were genotyped as described previously (van Rooijen et al., 2012) once cultured free from MEFs to avoid contamination of WT alleles. EpiSCs were cultured MEF-free culture on surface coated with fibronectin (Thermo Fisher Scientific) in medium containing 48% DMEM/F12, 48% Neurobasal Medium 1% B27 Supplement, 0.5% N2 Supplement B (all Thermo Fisher Scientific), 1x penicillin/streptomycin (100 units/mL), Fgf-basic (12 ng/mL), Activin A (20 ng/mL), IWP2 (2 μ M), 0.01% β -mercaptoethanol. Medium was refreshed on a daily basis and cells were passaged every two days. Cells cultures were routinely checked for mycoplasma contamination.

Induction was performed by withdrawal of IWP2 and addition of 3 μ M Chiron or 120 ng/ml Wnt3a (Cell Guidance Systems) or 3 μ M Chiron and 120 ng/ml Fgf8 (R&D Systems).

Histology and immunostaining

For histological analysis, tissues were fixed with 4% paraformaldehyde (PFA) overnight at 4°C. Embryos were imbedded in plastic (GMA Technovit type 8100) and sectioned at 7 μ m.

For immunofluorescence staining, 70 μ m-thick vibratome sections were made from embryos embedded in 4% low melting point agarose. EpiSCs were fixed with 4% PFA for 15 minutes at room temperature. Antibodies used were rabbit anti-Sox2 (Millipore), goat anti-*T Bra* (Santa Cruz, SC-17745) and mouse anti-*Cdx2* (Biogenex). DAPI was used for nuclear

staining.

LacZ reporter assays

Cdx2 and *T Bra* overlapping regions of interest were cloned upstream of a minimal *Hsp68* promoter using the Gateway Recombination System (Thermo Fisher Scientific). Primers used were: *Fgf8* proximal Fwd 5'-CTGCCTTTTGGCTGAGGAGA-3' and Rev 5'-GGGCTCTGGG-TAGAGTGTA-3'. *Fgf8* distal Fwd 5'-CACCCAATTTTCACGGTGGT-3' and Rev 5'-ACA-CACTAACGGTCTCAGCA-3'. *Rspo3* dist-2 Fwd 5'-TCTTCCCAAGTACTTTACAGCCT-3' and Rev 5'-TCAAACCTCATCATCCCCACA-3'. *Fgf4* Fwd 5'-TTGAGGCTGTGATCCCTG-TG-3' and Rev 5'-TGAGGGATGGGAAAGTCTGC-3'. *Fzd10* Fwd 5'-CGAGCACATCTT-TCGCCAAA-3' and Rev 5'-TTGCATCAGAAAGCCTTGCG-3'.

The *lacZ* transcription unit was isolated for micro-injection (Ittner and Gotz, 2007). Digested plasmid DNA was purified on Ultrafree-MC columns (Merck Millipore) and diluted to 2 ng/ μ L in micro-injection buffer (10 mM Tris pH 7.5, 0.1 mM EDTA). Embryos were harvested at E9.5 and fixed in 1% formaldehyde and 0.2% glutaraldehyde for 30 minutes on ice. Embryos were stained in a PBS buffer containing X-gal (1 mg/mL, ThermoFisher Scientific), potassium ferricyanide (5 mM), potassium ferrocyanide (5 mM), 2 mM magnesium chloride and 0.02% IGEPAL for up to 16 hours. Embryos were post-fixed in 4% PFA and stored in PBS. For each construct at least $n > 3$ positively stained embryos were analyzed. Embryos were genotyped using the primers: *lacZ* Fwd 5'-ATCCTCTGCATGGTCAGGTC-3' and Rev 5'-CGTGGCCTGATTCATTCC-3'.

Quantitative RT PCR

EpiSCs were washed with PBS and harvested in Trizol. RNA was isolated by Trizol (ThermoFisher Scientific) extraction followed by on-column (RNAeasy Purification Kit, Qiagen) DNA digestion by DNase I (Promega). 1 μ g total RNA was used as input for a reverse transcription reaction by SuperScriptII (ThermoFisher Scientific) using Oligo dTs (Promega) according to the manufacturer's protocol.

qPCR was performed on CFX Connect Real-Time PCR Detection System (Bio-Rad) using 2 or 3 μ L 1:10 diluted cDNA, primers (50 nM end concentration) and 2x SYBR Green Master Mix (Bio-Rad) in a 25 μ L total reaction. The $2^{-\Delta\Delta Ct}$ method was used for analysis. Primers used were: *Cdx2* Fwd 5'-TCCCTAGGAAGCCAAGTGA-3' and Rev 5'-AG-TGAAACTCCTTCTCCAGCTC-3'. *T Bra* Fwd 5'-GTATTCCCAATGGGGGTGGCT-3' and Rev 5'-CCTTAGAGCTGGGTACCTCTC-3'. *Cyp26a1* Fwd 5'-TGCTTCAGCGGAG-GAAGTTT-3' and Rev 5'-ACATTATCCGCGCCCATCAC-3'. *Wnt3a* Fwd 5'-ATGGTCTCT-CGGGAGTTTGC-3' and Rev 5'-GCACTTGAGGTGCATGTGAC-3'. *Fzd10* Fwd 5'-GCG-GATGATATAGCCCACCG-3' and Rev 5'-CCTTCCTCATCGACCCATCG-3'. *Rspo3* Fwd 5'-CCTTGAAAGTGCCTTGACAG-3' and Rev 5'-ACTCCATTCAGTGGCTCACA-3'. *Ppia* Fwd 5'-GGCCGATGACGAGCCC-3' and Rev 5'-TGTCTTTGGAACCTTGTCTGC-3'.

ChIP-seq and ChIP-qPCR

ChIP-seq was performed on 24h Chiron and Fgf8-induced WT EpiSCs cross-linked (1% formaldehyde) at RT for 15 min (TFs) or 10 min (histone marks). Chromatin was sheared to an average fragment size of 300 bp. Immunoprecipitations were performed overnight with 3 to 5 μ g of antibody. Antibodies used in ChIP-seq experiments were mouse anti-Cdx2 (Bio-genex, CDX2-88), goat anti-T Bra (Santa Cruz, sc-17745 X) and rabbit-anti H3K27ac (Abcam, ab4729). 10 ng of DNA was used for library preparation using the TruSeq DNA sample preparation kit (Illumina). Samples were run on an Illumina NextSeq500 sequencer by the Utrecht Sequencing Facility (USF; Utrecht, the Netherlands).

Enrichment of IP material was validated in E9.5 embryonic tissue by SYBR green qPCR and percent input or fold enrichment was calculated for at least two duplicate samples. Primers used were: *Hoxb4* Fwd 5'-TCCCCTTATTCAGCCTCCA-3' and Rev 5'-GGCAAT-TCCGTTTACAACCTGA-3'. *Hoxa5* Fwd 5'-CTGCCTTTTGAGGGTCTGT-3' and Rev 5'-CAGCCTGTGAGTCTGAAGCA-3'. *Hoxc5* Fwd 5'-CGCCCAAATTTACGACGACA-3' and Rev 5'-CCTCACCTTCAGATCCCTC-3'. *Fgf8* proximal Fwd 5'-CCCAGTCCTCCT-CCCTTTC-3' and Rev 5'-GGTTTTACAGCCCACTCAAGC-3'. *Fgf8* distal Fwd 5'-CC-CATTAGACTTTGCTGGGC-3' and Rev 5'-GTCGGTTGGAAAAGGGACAC-3'. *Rspo3* dist-2 Fwd 5'-GGTTTGGAGAGGCTGGAAGA-3' and Rev 5'-GGAGATGAAGCAAAG-CAGACA-3'. *Fgf4* Fwd 5'-GACTAGTCCAGTCCCAAC-3' and Rev 5'-CTAAAGCACTT-GAGGCCACC-3'. *Wnt5a* Fwd 5'-GCTGCTCTTGACTCTGAAGC-3' and Rev 5'-CCTGGG-TTTATGGTGGGTGT-3'. *Fzd10* Fwd 5'-CGCAAATCCCGTGTCTCTTTG-3' and Rev 5'-GTGGGAGGCCTTGAGATTGT-3'. Neg Chr17 primers from (Amin et al., 2015). Neg Chr3 primers from (Nishiyama et al., 2009).

Bioinformatic analysis

For ChIP-seq analysis, sequence reads were mapped to the mm9 reference genome using Bowtie v1.1.0 (Langmead et al., 2009). Reads that mapped to multiple locations and duplicate reads were removed with Samtools (Li et al., 2009). Peak calling was performed using MACS v2.1.0 (Zhang et al., 2008), using the matched input DNA reads as control. For Cdx2 and T Bra binding regions, peak calling was performed on pooled replicates with a similar Fripp (fraction of reads in peaks) score, with default settings and p-value 1e-5. Peak calling on H3K27ac samples was done with p-value 1e-5, extsize 300, llocal 100000 and nomodel parameters.

For ATAC-seq analysis, reads were trimmed to remove Nextera adapters and mapped to mm9 with Bowtie2 (Langmead and Salzberg, 2012) with -X 2000 parameter. Reads with MAPQ > 1, that mapped to chrM and duplicate reads were removed. Peak calling was done with llocal 100000 and nomodel parameters. For all bigWig tracks, reads were read per million normalized. Differentially enriched regions were identified using DESeq2 (Love et al., 2014) and descriptive analysis was performed for each individual sample.

CEL-seq processing was performed as described previously (Simmini et al., 2014)

Differential gene expression and statistical analysis was done using DESeq2 (Love et al., 2014). GO analysis was performed using GREAT v3.0 (McLean et al., 2010) and DAVID v6.7 (Huang da et al., 2009a, b). Motif analysis was performed using SeqPos on Galaxy/Cistrome (Liu et al., 2011).

Supplemental references

- Amin, S., Donaldson, I.J., Zannino, D.A., Hensman, J., Rattray, M., Losa, M., Spitz, F., Ladam, F., Sagerstrom, C., and Bobola, N. (2015). *Hoxa2* selectively enhances Meis binding to change a branchial arch ground state. *Developmental cell* 32, 265-277.
- Herrmann, B.G., Labeit, S., Poustka, A., King, T.R., and Lehrach, H. (1990). Cloning of the T gene required in mesoderm formation in the mouse. *Nature* 343, 617-622.
- Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic acids research* 37, 1-13.
- Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* 4, 44-57.
- Ittner, L.M., and Gotz, J. (2007). Pronuclear injection for the production of transgenic mice. *Nature protocols* 2, 1206-1215.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods* 9, 357-359.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* 10, R25.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078-2079.
- Liu, T., Ortiz, J.A., Taing, L., Meyer, C.A., Lee, B., Zhang, Y., Shin, H., Wong, S.S., Ma, J., Lei, Y., et al. (2011). Cistrome: an integrative platform for transcriptional regulation studies. *Genome biology* 12, R83.
- Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* 15, 550.
- McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nature biotechnology* 28, 495-501.
- Nishiyama, A., Xin, L., Sharov, A.A., Thomas, M., Mowrer, G., Meyers, E., Piao, Y., Mehta, S., Yee, S., Nakatake, Y., et al. (2009). Uncovering early response of gene regulatory networks in ESCs by systematic induction of transcription factors. *Cell stem cell* 5, 420-433.
- Simmini, S., Bialecka, M., Huch, M., Kester, L., van de Wetering, M., Sato, T., Beck, F., van Oudenaarden, A., Clevers, H., and Deschamps, J. (2014). Transformation of intestinal stem cells into gastric stem cells on loss of transcription factor *Cdx2*. *Nature communications* 5, 5728.
- Stringer, E.J., Duluc, I., Saandi, T., Davidson, I., Bialecka, M., Sato, T., Barker, N., Clevers, H., Pritchard, C.A., Winton, D.J., et al. (2012). *Cdx2* determines the fate of postnatal intestinal endoderm. *Development* 139, 465-474.
- van Rooijen, C., Simmini, S., Bialecka, M., Neijts, R., van de Ven, C., Beck, F., and Deschamps, J. (2012). Evolutionarily conserved requirement of *Cdx* for post-occipital tissue emergence. *Development* 139, 2576-2583.
- Zhang, Y., Liu, T., Meyer, C.A., Eeckhoutte, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome biology* 9, R137.

Chapter 4



Polarized regulatory landscape and Wnt responsiveness underlie Hox activation in embryos

Roel Neijts¹, Shilu Amin¹, Carina van Rooijen¹, Sander Tan¹, Menno P. Creyghton¹, Wouter de Laat¹ and Jacqueline Deschamps¹

¹Hubrecht Institute, Developmental Biology and Stem Cell Research, Uppsalalaan 8, 3584 CT Utrecht, and UMC Utrecht

Neijts R, Amin S, van Rooijen C, Tan S, Creyghton MP, de Laat W and Deschamps J. (2016) Polarized regulatory landscape and Wnt responsiveness underlie Hox activation in embryos, *Genes&Development* 30(17):1937–1942

Abstract

Sequential 3' to 5' activation of the Hox gene clusters in early embryos is a most fascinating issue in developmental biology. Neither the trigger nor the regulatory elements involved in the transcriptional initiation of the 3'-most Hox genes have been unraveled in any organism. We demonstrate that a series of enhancers, some of which Wnt-dependent, is located within a HoxA 3' subTAD. This subTAD forms the structural basis for multiple layers of 3'-polarized features, including DNA accessibility and enhancer activation. Deletion of the cassette of Wnt-dependent enhancers proves its crucial role in initial transcription of HoxA at the 3' side of the cluster.

Introduction

Correctly timed initiation of Hox transcription is fundamental to mediating the generation and patterning of nascent axial embryonic tissues (Kmita and Duboule, 2003; Deschamps and van Nes, 2005). Initial transcription of the earliest Hox genes in mouse embryos takes place in the posterior epiblast, in precursors of the extraembryonic and embryonic mesoderm at the late primitive streak stage (E7.2) during gastrulation (Forlani et al., 2003) (Fig. 1A). A striking feature is that the very first Hox gene is poised for transcription at the beginning of gastrulation, one full day before its transcription really starts (Forlani et al., 2003). Here we set out to investigate the molecular genetic interactions implementing the earliest Hox gene transcription.

Results and Discussion

Wnt activates Hox in a temporally colinear way

We reasoned that the trigger for Hox activation might be a canonical Wnt signal: *Wnt3* is present in the early posterior epiblast just before primitive streak formation (Rivera-Perez and Magnuson, 2005) and *Wnt3* null mutants do not express Hox genes (Liu et al., 1999). Pre-gastrulation embryos (E6.0, 'pre-streak', before primitive streak appearance) do not express Hox genes but exposure to CHIR99021 (Chiron, a Wnt agonist) for 10 hours strongly induces expression of the 3' paralogs *Hoxa1* and *Hoxb1* (Fig. 1B, Supplemental Fig. S1A). This observation led us to turn to a cell culture system very close to the pre-gastrulation epiblast: epiblast stem cells (EpiSCs) (Brons et al., 2007; Tesar et al., 2007). Activation of the Wnt pathway confers these cells a primitive streak-like identity (Kojima et al., 2014; Tsakiridis et al., 2015). We generated EpiSCs from wild type and *Wnt3* null pre-gastrula epiblasts and we cultured them in the presence of Wnt inhibitor IWP2 (Fig. 1C). Upon Wnt stimulation by Chiron, Hox genes were induced rapidly, reminiscent of what occurs in pre-streak embryos. Not only did Wnt activation induce the 3' most Hox gene, it induced the other Hox genes as well in a temporally colinear way, mimicking the *in vivo* situation (Izpisua-Belmonte et al., 1991; Deschamps and van Nes, 2005) (Fig. 1D and Supplemental Figs. S1B, S1C, Supplemental Table S1). We observed similar results upon inducing wild type and *Wnt3* null EpiSCs with Chiron (shown for *Hoxa1* in Supplemental Fig. S1D).

We compared the Hox chromatin states prior to and after Wnt exposure in EpiSCs that never experienced Wnt signaling before. We observed a dense coverage across the entire clusters by the repressive chromatin mark H3K27me3 – deposited by Polycomb complex 2 (PRC2) – and by PRC1 component Ring1b (Fig. 1E and Supplemental Fig. S2). Wnt activation leads to the removal of these repressive decorations, starting on the 3' side of the clusters and reaching completion around 72 hours after Chiron addition. The progressive loss of the H3K27me3 and Ring1b coverage was accompanied by the deposition of the activating marks H3K4me3 and H3K27ac, similarly to what was shown for *HoxD* in mid-gestation embryos (Soshnikova and Duboule, 2009) (Fig. 1E and Supplemental Fig. S2).

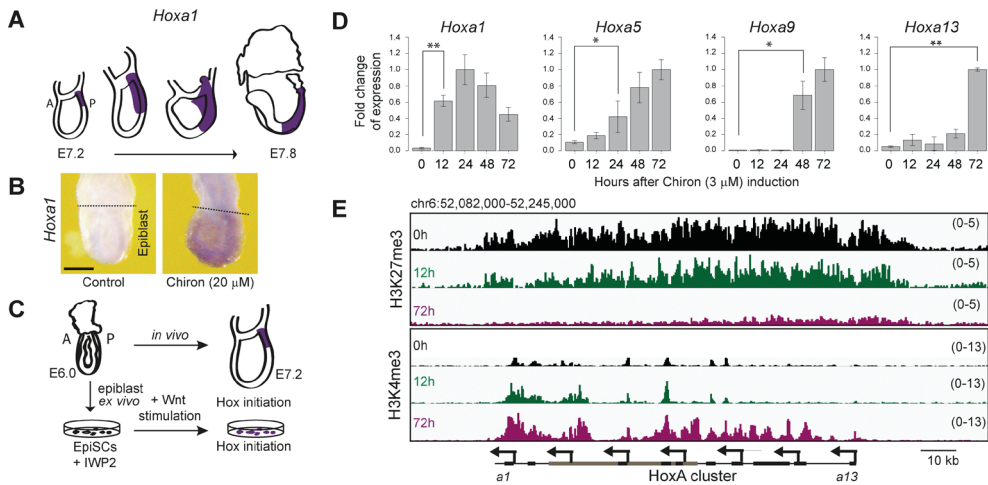


Figure 1: Transcriptional initiation of *HoxA* in early embryos and in EpiSCs. (A) Early expression pattern of *Hoxa1* in gastrulating embryos. A, anterior; P, posterior. (B) Precocious induction of *Hoxa1* expression in a E6.0 embryo by Chiron (10 hours). Scale bar = 100 μ m. (C) Parallelism between induction of *Hoxa1* in embryos and Wnt-stimulated induction of *Hoxa1* in EpiSCs. IWP2, Wnt inhibitor. (D) Kinetics of induction of HoxA genes by Chiron in wild type EpiSCs. Transcription measured by RT qPCR is relative to highest value of expression. Error bars, \pm SD. Asterisk, $p < 0.05$. Double asterisk, $p < 0.01$. (E) H3K27me3 and H3K4me3 marks along the HoxA cluster in uninduced (0h) and Wnt-induced *Wnt3* null EpiSCs (12h and 72h).

Putative early *HoxA* enhancers are located in a 3' subTAD

A key component of developmental gene regulation is the interaction between promoters and their regulatory landscapes (de Laat and Duboule, 2013). For in-depth analysis of early Hox regulation we further focused on the HoxA cluster, which is located at the boundary between two Topologically Associating Domains (TADs) (Dixon et al., 2012) (Supplemental Fig. S3A, top row). Chromosome conformation capture (4C-seq) shows that most HoxA genes interact much more frequently with regions lying in the 3' TAD than with sequences in the 5' TAD. Only the posterior-most *Hoxa13* mainly contacts remote sequences in the 5' TAD (Supplemental Fig. S3A). We observed that a proximal 300 kb genomic interval contains the majority of the interactions and, together with the 3' side of the HoxA cluster, forms a 3' subTAD (Fig. 2A, Supplemental Fig. S3A). Quantification of the interactions reveals that the more 3' the HoxA gene, the more it contacts the 3' subTAD, with *Hoxa1* showing a majority of these interactions, in particular with the proximal part of the 3' subTAD (Supplemental Fig. S3B). Comparison of 4C-seq profiles between embryonic stem cells (ESCs) and EpiSCs revealed that the *Hoxa1* interaction domain is less restricted in ESCs (Supplemental Fig. S3C), indicating that the 3' subTAD conformation is reinforced specifically in EpiSCs. We identified several H3K27ac peaks in the 3' subTAD that correspond to previously identified β -catenin binding regions (Zhang et al., 2013). Several of them overlap with 4C-seq interaction regions with *Hoxa1* (Fig. 2A), and are now referred to as 'HoxA Developmental Early Side' (Ades) putative enhancers. Three Ades regions are already acetylated before Wnt exposure (*Ades3-4* [part of the *Halr1* lncRNA locus (De Kumar et al., 2015)], *Ades5* and *Ades6*) and two depend

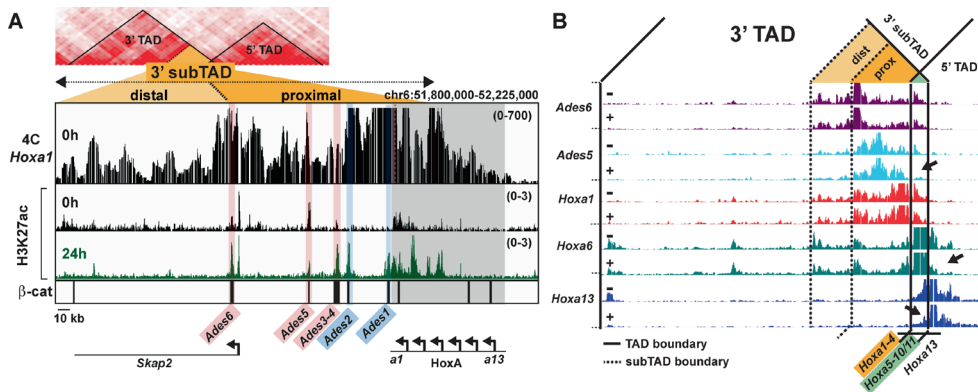


Figure 2: Interactions between the HoxA locus and putative Ades enhancers (A) Zoom in on 3' subTAD region with 4C-seq profile from *Hoxa1* viewpoint (red dotted line) and distribution of H3K27ac in uninduced (0h) and Chiron-induced (24h) EpiSCs. Positions acetylated before induction (*Ades3-4*, *Ades5* and *Ades6*) and positions becoming acetylated after induction (*Ades1* and *Ades2*) are highlighted in red and blue respectively. β -catenin (β -cat) binding regions (Zhang et al., 2013) are indicated. (B) 4C-seq profiles from *Ades* enhancers, and *HoxA* viewpoints in uninduced (-) and Wnt stimulated (+) conditions. The patterns of interactions define a proximal and a distal sub-part of the 3' subTAD. Wnt stimulation results in more compaction of these segments (arrows). *HoxA* appears to comprise three parts, indicated underneath. See also Supplemental Figs. S5 and S6.

on Wnt exposure (*Ades1* and *Ades2*) for their acetylation (Fig. 2A). *Ades1* lies in a region that was shown to drive part of the *Hoxa1* endogenous expression pattern (Frasch et al., 1995; Nolte et al., 2013). In reciprocal 4C-seq experiments, we found that the *Ades* sites show interactions with the 3' part of the Hox cluster (Supplemental Fig. S4). The 3' subTAD which we identified in EpiSCs is therefore a domain of intensive interactions between the cluster and its putative enhancers.

The HoxA landscape is segmented in several sub-domains with particular interaction specificities

We investigated whether chromosome architecture changes during Hox initiation in EpiSCs. Chromosomal interactions of *Hoxa1* with the *Ades* region within the 3' TAD do not change upon Wnt stimulation (Fig. 2B, Supplemental Fig. S5A), indicating a constitutively active conformation of this region. In uninduced conditions, contacts between the most proximal *Ades* enhancers (*Ades2*, *Ades3-4* and *Ades5*) and Hox genes involve 3' and middle Hox genes (Supplemental Figs. 4B, 5B). Chiron induction leads to a further 3' restriction of these contacts to the 3' subTAD (Supplemental Fig. S5B). Activation of HoxA transcription therefore leads to a further focalisation of the already confined enhancer-3' Hox interactions occurring within the 3' TAD. The situation is different for *Ades6* that lies at an internal boundary of the 3' subTAD and is less confined in its interactions, maintaining contacts with middle Hox genes on the one side and reaching the distal part of the 3' subTAD on the other side (Fig. 2B, Supplemental Fig. S5B). The 3' subTAD thus appears to be segmented into a proximal and a distal

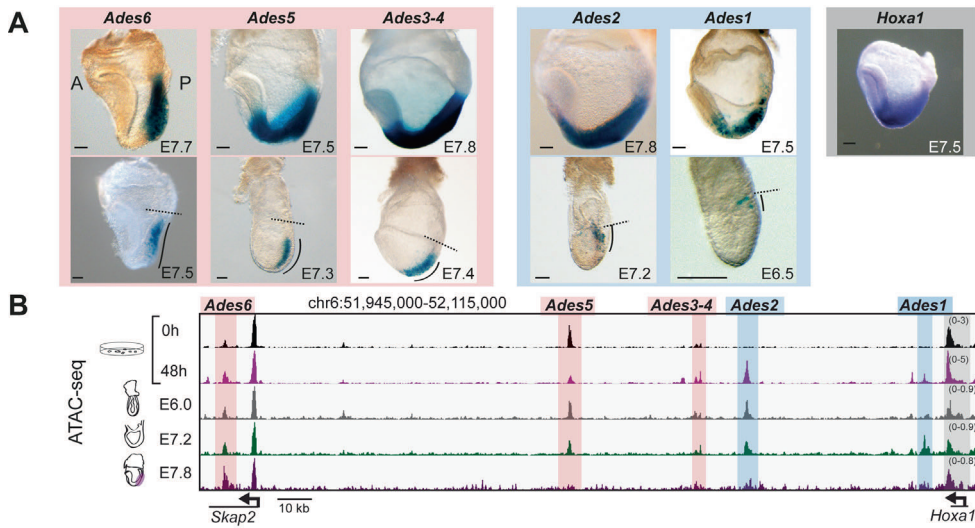


Figure 3: Activity and DNA accessibility of the Ades enhancers. (A) First row, activity of Ades enhancers coupled to lacZ in E7.5 to E7.8 (head fold to early somite) embryos. On the right, *Hoxa1* expression. Second row, the earliest embryonic stage in which each enhancer is observed to be active (varying from E6.5 – before actual *Hoxa1* expression – to E7.5). Black curved line, region of activity. Dotted line, boundary between embryonic (below the line) and extraembryonic tissues (above the line). (B) ATAC-seq profile of EpiSCs (uninduced and after 48 hours of Chiron activation) and in pre-Hox (E6.0), early Hox (E7.2) and later Hox (posterior tissues of E7.8) embryos in the Ades region. All scale bars are 100 μ m.

part (Fig. 2 and Supplemental Fig. S5B). Seen from the Hox viewpoints, a relative loosening of the interactions of middle HoxA genes (*Hoxa6* and *Hoxa10*) from their 5' surrounding, and a decrease in interactions of *Hoxa13* with its 3' surrounding is observed upon Wnt stimulation, demarcating a virtual boundary located between *Hoxa10* and *Hoxa13* (Fig. 2B, Supplemental Fig. S6). These data in uninduced and stimulated EpiSCs revealed that the HoxA locus and its 3' TAD are segmented in different *cis* domains, which become more compact during Wnt activation. According to the preference of interactions, the HoxA cluster itself appears to be subdivided in three parts: 3' genes (*Hoxa1-4*), middle HoxA genes on the overlap between the 3' and 5' TADs (*Hoxa5-10/11*), and *Hoxa13* that belongs to the 5' TAD (Figure 2B).

Differential activation of Ades enhancers during development

The biological activity of the Ades putative enhancers was tested *in vivo* using *lacZ* reporters. Each of the five Ades regions appeared to exhibit transcription enhancing activity that reproduces aspects of the spatiotemporal expression pattern of *Hoxa1* between E6.5 and mid-gestation (Fig. 3A and Supplemental Fig. S7). *Ades1* and *Ades2* are active the earliest, perfectly mimicking the temporal and spatial features of *Hoxa1* initial transcription in the posterior-most epiblast and extraembryonic mesoderm precursors of the nascent allantois (Fig. 3A, second row and Supplemental Fig. S7). *Ades3-4* and *Ades5* are active from more anterior posi-

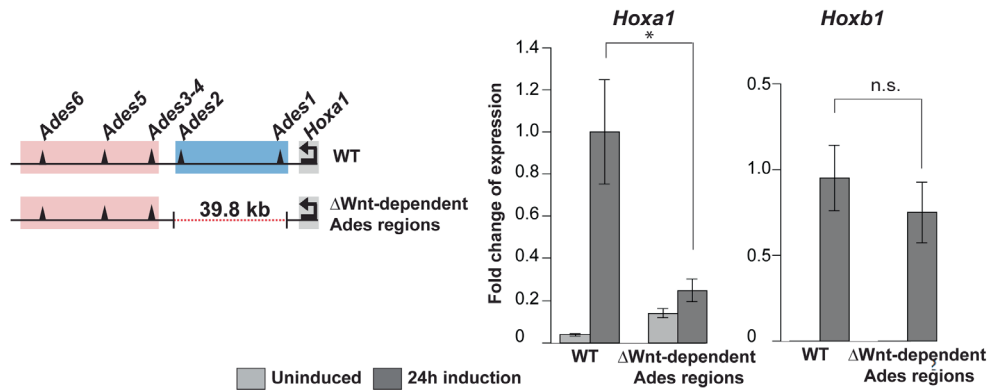


Figure 4: The most proximal region of the 3' subTAD is required to activate *Hoxa1*. Deletion of the Wnt-dependent *Ades1* and *Ades2* region reduces *Hoxa1* transcriptional response to Chiron (24 hours); *Hoxb1* is unaffected. Transcription measured by RT qPCR is relative to highest value of expression. Errors bars, \pm SD. Asterisk, $p < 0.05$.

tions of the primitive streak, excluding the posterior streak and extraembryonic mesoderm (Fig. 3A, second row). *Ades6* drives a pattern restricted to a subset of tissues expressing endogenous *Hoxa1*, starting at a later time point than the other *Ades* regions (Fig. 3A). *Ades1* and *Ades2*, which depend on Wnt for their activation, thus are the earliest active enhancers driving Hox expression.

To obtain information on the DNA accessibility of the different *Ades* enhancer chromatin *in vivo* and in EpiSCs, we performed Assay for Transposase-Accessible Chromatin (ATAC)-seq experiments (Buenrostro et al., 2013). *Ades3-4*, *Ades5* and *Ades6* are accessible in uninduced EpiSCs, whereas *Ades1* and *Ades2* are not yet opened in these conditions (Fig. 3B). Strikingly, *Ades1* and *Ades2* are accessible in Wnt-treated EpiSCs and in pre-streak (E6.0) and older embryos (E7.2 and posterior part of early somite E7.8) (Fig. 3B). Comparing the kinetics of chromatin opening of the strictly Wnt-dependent *Ades1* and *Ades2* enhancers in EpiSCs and embryos suggests that uninduced EpiSCs represent the naïve state of the Hox neighborhood ('pre-primed' Hox state), and that the pre-streak embryos, which already accumulated *Wnt3* in their posterior region (Rivera-Perez and Magnuson, 2005), are in a primed Hox state. These pre-streak embryos have their HoxA cluster open already (Supplemental Fig. S8A, left panel). In embryos and EpiSCs that have started Hox expression, all *Ades* enhancers (Fig. 3B) and HoxA genes (Supplemental Fig. S8A) are accessible. Interestingly, the 5' Hox TAD in E7.8 embryos displays accessible positions corresponding to limb-specific enhancers (Berlivet et al., 2013; Lonfat et al., 2014) (Supplemental Fig. 8A, right panel). These 5' regulatory regions never become open or acetylated upon Wnt exposure in EpiSCs, showing that the Wnt response is restricted to the 3' side of the HoxA locus and that the gradual transcriptional activation is independent of the 5' neighborhood (Supplemental Fig. S8B).

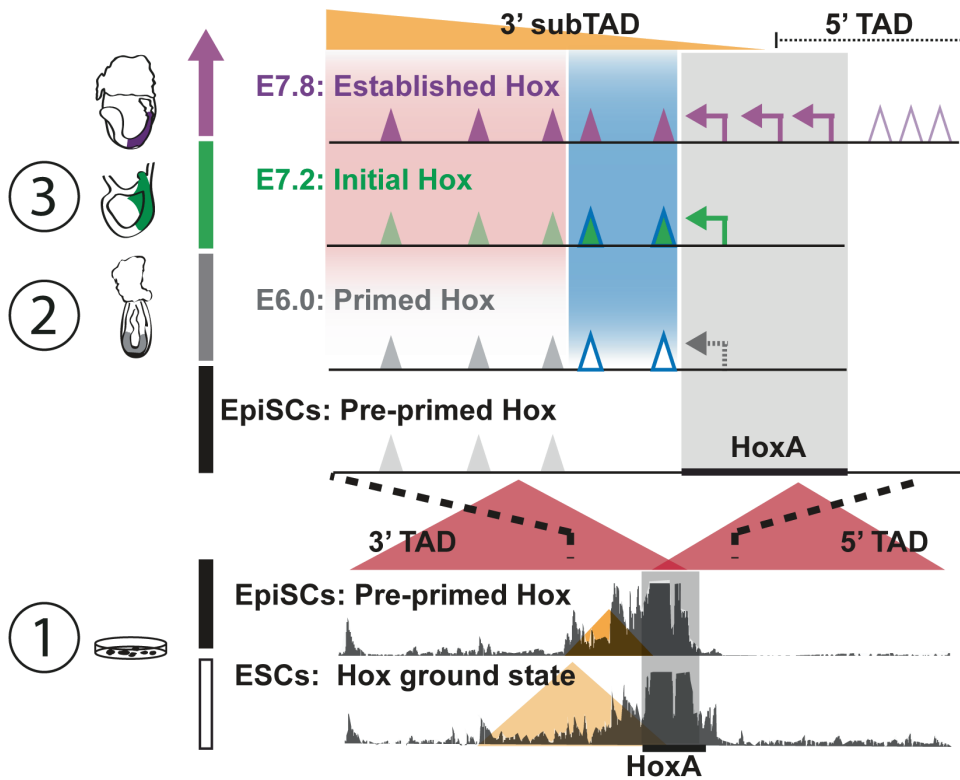


Figure 5: Model summarizing the findings of distinct steps leading to transcriptional initiation of HoxA genes. Three successive phases of 3' oriented epigenetic events culminate in 3' Hox gene transcription. (1) Tropism of contacts between 3' HoxA and the 3' surrounding in ESCs (Hox ground state) and its compaction to the 3' subTAD in EpiSCs (pre-primed Hox state). (2) Accessibility of Wnt-dependent proximal elements (open triangles) in the proximal 3' subTAD appears between uninduced EpiSCs and E6.0 embryos. (3) Acetylation of these enhancers and 3' HoxA transcription arise at E7.2. At that stage, all enhancers are acetylated. More 5' HoxA genes are subsequently expressed (E7.8).

A proximal region in the 3' subTAD is required for Wnt-induced Hox initiation

Our results obtained by H3K27ac ChIP-seq and ATAC-seq reveal that the HoxA 3' subTAD contains Wnt-responsive enhancers. The developmentally early active enhancers *Ades1* and *Ades2* that depend on Wnt for their activation are located in the proximal part of the *Ades* region, whereas the enhancers already acetylated before Wnt exposure reside more distally. To directly test the functional significance of the proximal Wnt-dependent region for Hox gene activation, we engineered its CRISPR/Cas9-directed deletion. We excised a 39.8 kb region proximal to *Hoxa1* in ESCs (Fig. 4, Supplemental Fig. S9A). After differentiation of the ESCs into EpiSCs and subsequent Wnt activation (Supplemental Fig. S9B), we measured transcript levels of *Hoxa1* and its paralog *Hoxb1* (Fig. 4). We observed that deletion of the Wnt-dependent genomic interval leads to severely reduced Wnt-responsiveness of *Hoxa1* transcription.

The proximal part of the HoxA 3' subTAD therefore is crucial for the transcriptional response of the 3' end of the HoxA cluster to incoming Wnt signals. Deletion of this region also caused a decrease in Wnt response of *Hoxa5*, whereas more 5' genes *Hoxa7* and *Hoxa9* remain highly inducible (Supplemental Fig. S9C). Deletion of the *Ades1-2* region thus compromises transcriptional activation of *Hoxa1* and affects the sensitivity of more 5' HoxA genes for Wnt signals. Some of the effects of the deletion on the expression of 5' Hox genes may result from the reduction in the distance between these genes and distal Ades enhancers.

In summary, we identified several layers of regulatory events that modify and activate the HoxA cluster neighborhood on its 3' side (Fig. 5). On top of a tropism of interactions of the earliest HoxA genes with their segmented 3' regulatory landscape, we have found a bipartite activating regulatory module. A Wnt-dependent proximal cassette of enhancers (*Ades1* and *Ades2*) responds to incoming Wnt signals in the posterior primitive streak region and initiates 3' HoxA transcription. A more distal cassette (*Ades3-4* to *Ades6*) consolidates gene expression via enhancers that are activated slightly later in development. These activating modules and early Wnt-dependence are exclusively located at the 3' side of the HoxA locus. We propose that it is the 3' restriction of consecutive chromatin opening and enhancer activation in this 3' HoxA region that dictates the first Hox transcription on the early side of the cluster, in response to the earliest Wnt signal in the gastrulating embryo. Our data elucidate the molecular genetics underlying one of the most intriguing and evolutionarily conserved developmental systems: the spatiotemporally controlled turning on of the Hox genes when the body plan is laid down during early embryogenesis.

Material and methods

Animals

All mice used are in the C57Bl6j/CBA mixed background. Heterozygous *Wnt3* mutants were generated by intercross of a *Wnt3* conditional strain (Barrow et al., 2003) with *Sox2Cre* mice (Hayashi et al., 2002). Ear clip DNA was used for genotyping (primers in Supplemental Table S2). All animal experiments were performed in accordance with institutional and national regulations, under control of the Central Animal Experiments Committee.

Embryo culture

E6.0 embryos were isolated in M2 medium (Sigma Aldrich) and cultured in 4-well plates (Nunc, ThermoFisher Scientific) for 10 hours in medium containing 75% Knock Out Serum Replacement (KOSR), 25% DMEM/F12, 1% non-essential amino acids (NEAAs), and 2 mM L-glutamine (all ThermoFisher) in the presence of 20 μ M CHIR99021 ('Chiron', Cell Guidance Systems, in DMSO) or DMSO only.

In situ hybridization (ISH)

ISH was performed as described elsewhere (Young et al., 2009). A *Hoxa1* probe was generated from a 1.2 kb insert transcribed by T7 polymerase (Promega). The *Hoxb1* probe was described elsewhere (Marshall et al., 1994). For each experiment $n > 3$ embryos were used.

EpiSC culture and Hox induction

E6.0 embryos were isolated in M2 medium; extraembryonic tissue and primitive endoderm were removed. Explants were cultured on mitotically inactivated mouse embryonic fibroblasts (MEFs) in medium containing 20% KOSR, 77% DMEM/F12, 1x penicillin/streptomycin (pen/strep) (100 units/mL), 2 mM L-glutamine, 1x NEAAs, 0.01% β -mercaptoethanol (β -ME), Fgf-basic (12 ng/mL, ThermoFisher), Activin A (20 ng/mL, R&D Systems) and Wnt inhibitor IWP2 (2 μ M in DMSO, Sigma Aldrich). *Wnt3* null EpiSCs were genotyped once cultured free of MEFs. MEF-free culture took place on surface coated by fibronectin (ThermoFisher) in medium containing 48% DMEM/F12, 48% Neurobasal Medium 1% B27 Supplement, 0.5% N2 Supplement B (all ThermoFisher), 1x pen/strep, Fgf-basic (12 ng/mL), Activin A (20 ng/mL), IWP2 (2 μ M), 0.01% β -ME. Medium was refreshed on a daily basis. Cells were tested for mycoplasma contamination. Wnt stimulation was performed after at least two passages feeder-free culturing by withdrawal of IWP2 and addition of Chiron (3 μ M).

RT qPCR

RNA was isolated by combining Trizol (ThermoFisher) extraction followed by on-column (RNAeasy Purification Kit, Qiagen) DNase I (Promega) treatment. 1 μ g total RNA was used for reverse transcription by SuperScriptII (ThermoFisher) using Oligo dTs (Promega). qPCR was performed on 1/10 (Hox genes) or 1/100 (reference gene *Ppia*) cDNA dilutions. Primers sequences are in Supplemental Table S3. The reference gene was verified to be expressed constantly and not to respond to Wnt. 2 or 3 μ L diluted cDNA, primers (50 nM) and SYBR Green Master Mix (Bio-Rad) were used in 20 μ L reactions. Measurements were done on CFX Connect Real-Time PCR Detection System (Bio-Rad). Two to three biological and three technical replicates were used. For analysis the $2^{-\Delta\Delta C_t}$ method (Livak and Schmittgen, 2001) was used. Comparison between samples was performed using Student's t test.

ChIP-seq

ChIP was performed on 10 million feeder-free EpiSCs. Cells were cross-linked (1% formaldehyde) for 10 minutes. Crosslinking was stopped by glycine (125 mM end concentration). Nuclei were isolated in lysis buffer 1 (50 mM HEPES, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% IGEPAL (Sigma Aldrich), 0.25% Triton X-100 (Sigma Aldrich), 1 mM PMSF and 1x PIC (Roche)) for 10 minutes on ice. Nuclei were lysed with lysis buffer 2 (1% SDS, 50 mM Tris pH 8.0, 10 mM EDTA, 1 mM PMSF and 1x PIC) on ice. Samples were sonicated on a Diagenode Bioruptor, for twice 20 cycles (30 seconds on/off, high power) at 4°C. Samples were rotated overnight at 4°C with 5 μ g of the appropriate antibody: H3K4me3 (Abcam, ab8580),

H3K27ac (Abcam, ab4729), H3K27me3 (Millipore, 17-622) or Ring1b (39663, ActiveMotif). The next day, pre-washed Dynabeads Protein G beads (ThermoFisher) were added. After 4 hours, beads were washed with HEPES buffer (20 mM HEPES, 1 mM EDTA, 0.5% IGEPAL, 150 mM NaCl, 1 mM PMSF and 1x PIC), with LiCl buffer (250 mM LiCl, 1% IGEPAL, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris pH 8.0, 1 mM PMSF and 1x PIC) and with TE containing 1 mM PMSF and 1x PIC at 4°C. The bound complexes were eluted in elution buffer (1% SDS, 100 mM NaHCO₃) and de-cross-linked overnight. The next day, RNase A (30 minutes at 37°C) and Proteinase K (3 hours at 55°C) treatments followed. DNA was extracted by phenol/chloroform. Concentration and quality was checked by Qubit (ThermoFisher) and Bioanalyzer (Agilent) respectively. 10 ng of DNA was used for TruSeq DNA sample preparation kit (Illumina). Samples were run by the MIT BioMicro Center (Cambridge, USA) or the Utrecht Sequencing Facility (USF) (Utrecht, the Netherlands) on an Illumina HiSeq2000.

4C-seq

4C-seq was performed on 5-10 million feeder-free EpiSCs or ESCs, according to published protocols (Splinter et al., 2012). A first digest was performed with DpnII (NEB); a second digest with Csp6I (ThermoFisher). PCR primers were designed using guidelines described previously (Splinter et al., 2012) and are listed in Supplemental Table S4. Samples were run by the USF on an Illumina NextSeq500. After mapping on a reduced mm9 genome (van der Werken et al., 2012), the highest covered fragment was removed and the dataset was normalized to 1 million intrachromosomal reads.

LacZ reporter assays

Ades enhancers were cloned upstream of a minimal *Hsp68* promoter (Pennacchio et al., 2006). Primers are in Supplemental Table S5. The vector was linearized for micro-injection. DNA concentration was filtered (0.45 µm column, Millipore), diluted to 2 ng/µL, re-filtered (0.22 µm column, Millipore) (Ittner and Gotz, 2007). Embryos were harvested between E6.5 and E9.5 and fixed in 1% formaldehyde and 0.2% glutaraldehyde. Staining took place in PBS containing X-gal (1 mg/mL, ThermoFisher), C₆N₆FeK₃ (5 mM), C₆N₆FeK₄ (5 mM), 2 mM MgCl₂ and 0.02% IGEPAL. For each construct n>5 positively stained embryos were analyzed. Embryos were genotyped by primers listed in Supplemental Table S2.

CRISPR/Cas9-directed genomic deletion

Mouse ESCs (129/Ola-derived IB10) were cultured on MEFs in G-MEM containing 10% FCS, 1x Glutamax, 1 mM sodium pyruvate, 1x NEAAs, 0.01% β-ME (all ThermoFisher) and LIF (10 ng/mL, Millipore). sgRNA sequences were designed using crispr.mit.edu (see Supplemental Table S6) and cloned into pX330 (Cong et al., 2013). A puromycin-resistance vector was co-transfected with the pX330 plasmids using Lipofectamine 3000 (ThermoFisher). Transfection took place in feeder-free conditions, in medium conditioned on Buffalo Rat Liver (BRL) cells (60%), supplemented with LIF and β-ME. Cells were selected for 48 hours with puromy-

cin (2 µg/mL). ESCs were re-plated and after 3 to 4 days at least 24 colonies were picked. Cells were passaged, frozen and genotyped. Primers are in Supplemental Table S2. Positive clones were cultured and differentiated to EpiSCs by culturing them in EpiSC medium (including IWP2) for 3 weeks on MEFs.

ATAC-seq

ATAC-seq was performed according to the standard protocol (Buenrostro et al., 2013) on 50,000 EpiSCs and on embryos (E6.0, E7.2, posterior parts of E7.8). Embryos were treated with collagenase for 30 minutes at 37 °C and cells were suspended by mild pipetting. Number of embryos used per assay: 25 (E6.0), 5 (E7.2), 7 posterior parts (E7.8). Nuclei were lysed, subjected to Tn5 transposase (Illumina), and DNA was isolated by MinElute columns (Qia-gen). After PCR (maximum of 9 cycles), tagmented DNA was purified using AmpureBeads (1.6 times volume) (Betancourt). Concentration and quality was checked by Qubit and Bio-analyzer respectively. Samples were paired-end sequenced (Illumina NextSeq500) by the USE.

All sequencing data is available under GEO Accession Number GSE81203.

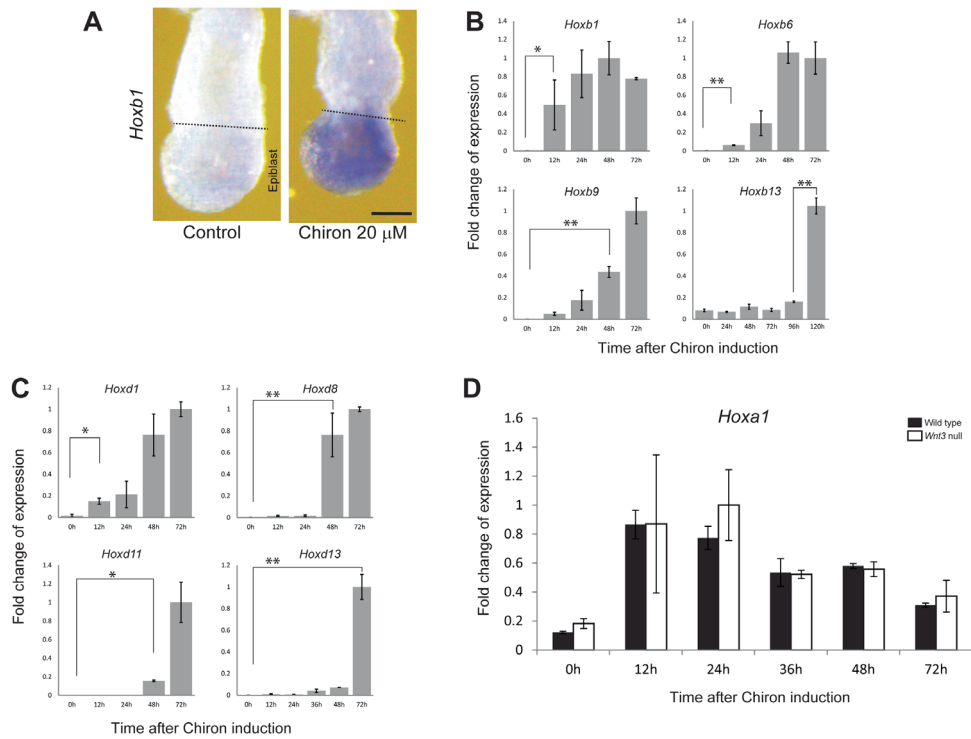
Acknowledgments

We thank S. van den Brink for help with ESCs and J. Korving for micro-injection. We are thankful to C. Vermeulen, G. Geeven and A. Melquiond (laboratory W.d.L.) for their help on 4C-seq. We are grateful to M. Vermunt and C. Wiggers (laboratory M.P.C.) for their advice on ChIP-seq and ATAC-seq. T. Young generated the *Hoxa1* probe. We thank D. Duboule and J. Zakany for providing the pX330 vector. We thank S. Simmini for reading of the manuscript. This work was supported by the Netherlands Institute for Regenerative Medicine (NIRM, grant FES0908). J.D. and R.N. conceived the study. R.N., S.A. and C.v.R. performed experiments. S.T. performed analyses on ChIP-seq and ATAC-seq. M.P.C. and W.d.L. provided expert advice on parts of the experiments. R.N. and J.D. wrote the manuscript.

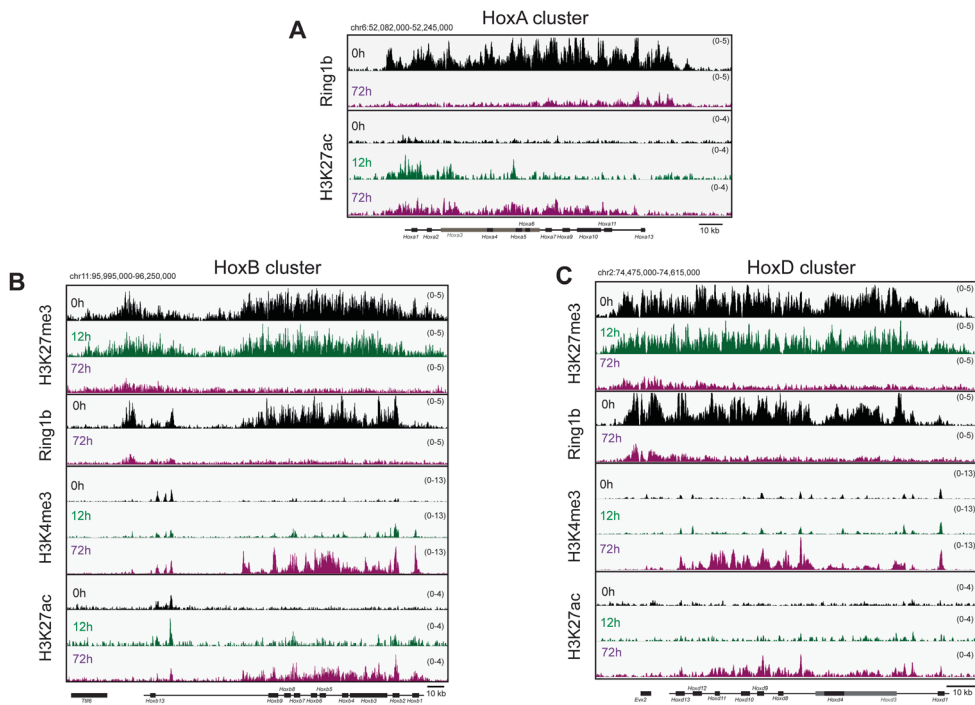
References

- Barrow JR, Thomas KR, Boussadia-Zahui O, Moore R, Kemler R, Capecchi MR, McMahon AP. 2003. Ectodermal Wnt3/beta-catenin signaling is required for the establishment and maintenance of the apical ectodermal ridge. *Genes Dev* 17:394-409.
- Berlivet S, Paquette D, Dumouchel A, Langlais D, Dostie J, Kmita M. 2013. Clustering of tissue-specific sub-TADs accompanies the regulation of HoxA genes in developing limbs. *PLoS Genet* 9:e1004018.
- Brons IG, Smithers LE, Trotter MW, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett SK, Clarkson A, Ahrlund-Richter L, Pedersen RA, Vallier L. 2007. Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* 448:191-195.
- Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* 10:1213-1218.
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339:819-823.
- De Kumar B, Parrish ME, Slaughter BD, Unruh JR, Gogol M, Seidel C, Paulson A, Li H, Gaudenz K, Peak A, McDowell W, Fleharty B, Ahn Y, Lin C, Smith E, Shilatifard A, Krumlauf R. 2015. Analysis of dynamic changes in retinoid-induced transcription and epigenetic profiles of murine Hox clusters in ES cells. *Genome Res* 25:1229-1243.
- de Laat W, Duboule D. 2013. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* 502:499-506.
- Deschamps J, van Nes J. 2005. Developmental regulation of the Hox genes during axial morphogenesis in the mouse. *Development* 132:2931-2942.
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485:376-380.
- Forlani S, Lawson KA, Deschamps J. 2003. Acquisition of Hox codes during gastrulation and axial elongation in the mouse embryo. *Development* 130:3807-3819.
- Frasch M, Chen X, Lufkin T. 1995. Evolutionary-conserved enhancers direct region-specific expression of the murine Hoxa-1 and Hoxa-2 loci in both mice and *Drosophila*. *Development* 121:957-974.
- Hayashi S, Lewis P, Pevny L, McMahon AP. 2002. Efficient gene modulation in mouse epiblast using a Sox2Cre transgenic mouse strain. *Mech Dev* 119 Suppl 1:S97-S101.
- Ittner LM, Gotz J. 2007. Pronuclear injection for the production of transgenic mice. *Nat Protoc* 2:1206-1215.
- Izpisua-Belmonte JC, Falkenstein H, Dolle P, Renucci A, Duboule D. 1991. Murine genes related to the *Drosophila* AbdB homeotic genes are sequentially expressed during development of the posterior part of the body. *EMBO J* 10:2279-2289.
- Kmita M, Duboule D. 2003. Organizing axes in time and space; 25 years of colinear tinkering. *Science* 301:331-333.
- Kojima Y, Kaufman-Francis K, Studdert JB, Steiner KA, Power MD, Loebel DA, Jones V, Hor A, de Alencastro G, Logan GJ, Teber ET, Tam OH, Stutz MD, Alexander IE, Pickett HA, Tam PP. 2014. The transcriptional and functional properties of mouse epiblast stem cells resemble the anterior primitive streak. *Cell Stem Cell* 14:107-120.
- Liu P, Wakamiya M, Shea MJ, Albrecht U, Behringer RR, Bradley A. 1999. Requirement for Wnt3 in vertebrate axis formation. *Nat Genet* 22:361-365.
- Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25:402-408.
- Lonfat N, Montavon T, Darbellay F, Gitto S, Duboule D. 2014. Convergent evolution of complex regulatory landscapes and pleiotropy at Hox loci. *Science* 346:1004-1006.
- Marshall H, Studer M, Popperl H, Aparicio S, Kuroiwa A, Brenner S, Krumlauf R. 1994. A conserved retinoic acid response element required for early expression of the homeobox gene Hoxb-1. *Nature* 370:567-571.
- Nolte C, Jinks T, Wang X, Martinez Pastor MT, Krumlauf R. 2013. Shadow enhancers flanking the HoxB cluster direct

- dynamic Hox expression in early heart and endoderm development. *Dev Biol* 383:158-173.
- Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, Shoukry M, Minovitsky S, Dubchak I, Holt A, Lewis KD, Plajzer-Frick I, Akiyama J, De Val S, Afzal V, Black BL, Couronne O, Eisen MB, Visel A, Rubin EM. 2006. In vivo enhancer analysis of human conserved non-coding sequences. *Nature* 444:499-502.
- Rivera-Perez JA, Magnuson T. 2005. Primitive streak formation in mice is preceded by localized activation of Brachyury and Wnt3. *Dev Biol* 288:363-371.
- Soshnikova N, Duboule D. 2009. Epigenetic temporal control of mouse Hox genes in vivo. *Science* 324:1320-1323.
- Splinter E, de Wit E, van de Werken HJ, Klous P, de Laat W. 2012. Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: from fixation to computation. *Methods* 58:221-230.
- Tesar PJ, Chenoweth JG, Brook FA, Davies TJ, Evans EP, Mack DL, Gardner RL, McKay RD. 2007. New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* 448:196-199.
- Tsakiridis A, Huang Y, Blin G, Skylaki S, Wymeersch F, Osorno R, Economou C, Karagianni E, Zhao S, Lowell S, Wilson V. 2015. Distinct Wnt-driven primitive streak-like populations reflect in vivo lineage precursors. *Development* 142:809.
- van de Werken HJ, Landan G, Holwerda SJ, Hoichman M, Klous P, Chachik R, Splinter E, Valdes-Quezada C, Oz Y, Bouwman BA, Verstegen MJ, de Wit E, Tanay A, de Laat W. 2012. Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat Methods* 9:969-972.
- Young T, Rowland JE, van de Ven C, Bialecka M, Novoa A, Carapuco M, van Nes J, de Graaff W, Duluc I, Freund JN, Beck F, Mallo M, Deschamps J. 2009. Cdx and Hox genes differentially regulate posterior axial growth in mammalian embryos. *Dev Cell* 17:516-526.
- Zhang X, Peterson KA, Liu XS, McMahon AP, Ohba S. 2013. Gene regulatory networks mediating canonical Wnt signal-directed control of pluripotency and differentiation in embryo stem cells. *Stem Cells* 31:2667-2679.



Supplemental Figure S1: Wnt-triggered transcriptional induction of Hox genes *in vivo* and in EpiSCs. (A) Precocious induction of *Hoxb1* by the Wnt agonist CHIR99021 ('Chiron', 20 μ M) for 10 hours in pre-gastrulation (pre-streak) E6.0 embryos. The epiblast is indicated. Scale bar is 100 μ m. (B) Induction of HoxB genes *Hoxb1*, *Hoxb6*, *Hoxb9* and *Hoxb13* by Chiron (3 μ M) in wild type EpiSCs. (C) Kinetics of induction of HoxD genes *Hoxd1*, *Hoxd8*, *Hoxd11* and *Hoxd13* by Chiron (3 μ M) in wild type EpiSCs. (D) Comparison of expression dynamics of *Hoxa1* between wild type and *Wnt3* null EpiSCs, treated with Chiron (3 μ M). Over the entire time course transcription levels are comparable. Errors bars represent \pm SD. Asterisk indicates $p > 0.05$. Double asterisk indicates $p > 0.01$. Transcription over time was measured by qRT-PCR and calculated relative to the highest value of expression. Note, *Hoxc* is not included in this study since it lacks the 3' part of the cluster (its 3'-most gene is *Hoxc4*).



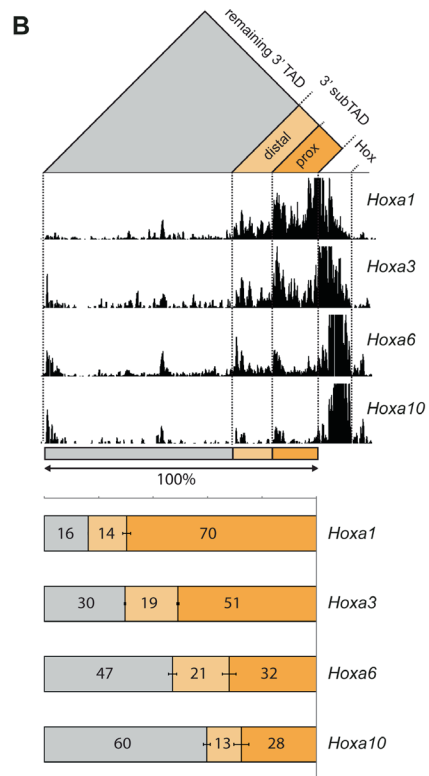
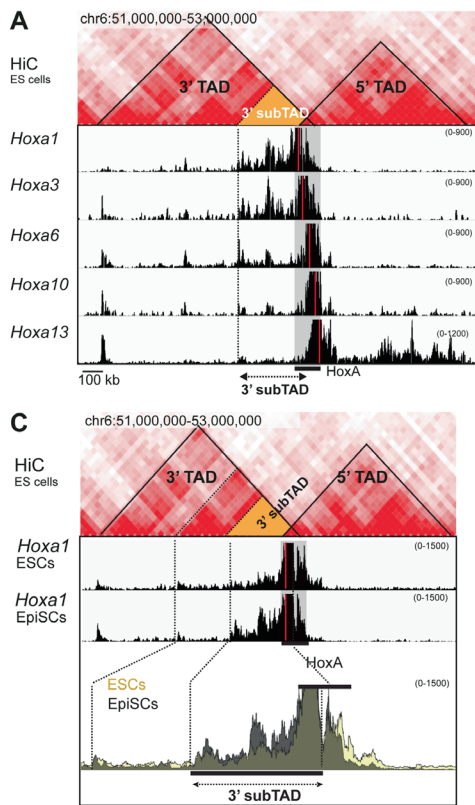
Supplemental Figure S2: Dynamics of chromatin mark coverage during Wnt-triggered induction of Hox genes in EpiSCs. (A) Distribution of Ring1b and H3K27ac histone mark along the HoxA cluster in uninduced *Wnt3* null EpiSCs (black profiles) and *Wnt3* null EpiSCs induced for 12 and 72 hours (green and purple profiles respectively) with Chiron (3 μ M). (B, C) Distribution of H3K27me3, H3K4me3 and H3K27ac decoration and Ring1b binding along the HoxB (B) and HoxD (C) clusters (colors as in A). Note, HoxC is not included in this study since it lacks the paralogous 3' part of the cluster (the 3'-most gene is *Hoxc4*).

(Right page, on top)

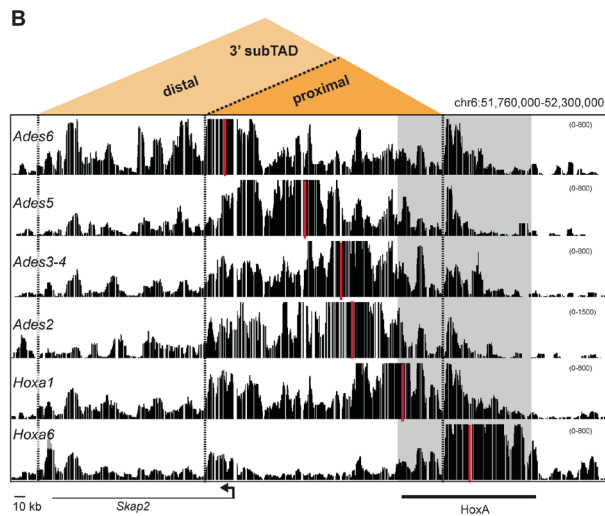
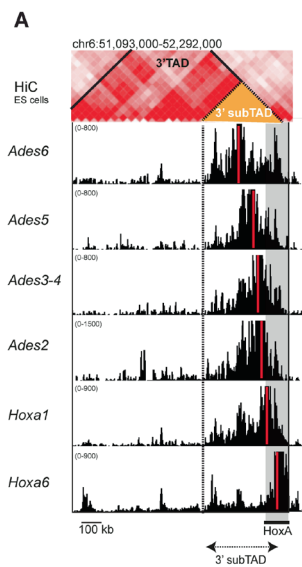
Supplemental Figure S3: Genomic organization and definition of the enhancer-containing 3' subTAD. (A) TAD organization in the HoxA neighborhood (Dixon et al., 2012) and 4C-seq profiles throughout the HoxA cluster in uninduced EpiSCs. A 300 kb region proximal to HoxA is indicated as 3' subTAD (B) Quantification of interactions of HoxA viewpoints with the proximal and distal parts of the 3' subTAD, relative to the number of all interactions in the entire 3' TAD (100%). Interactions within the HoxA cluster are excluded. (C) 4C-seq profiles of interactions of *Hoxa1* in ESCs and EpiSCs. Lower panel, comparison of these profiles. The overlap between yellow and black is shown as dark grey. Red lines indicate 4C-seq viewpoints.

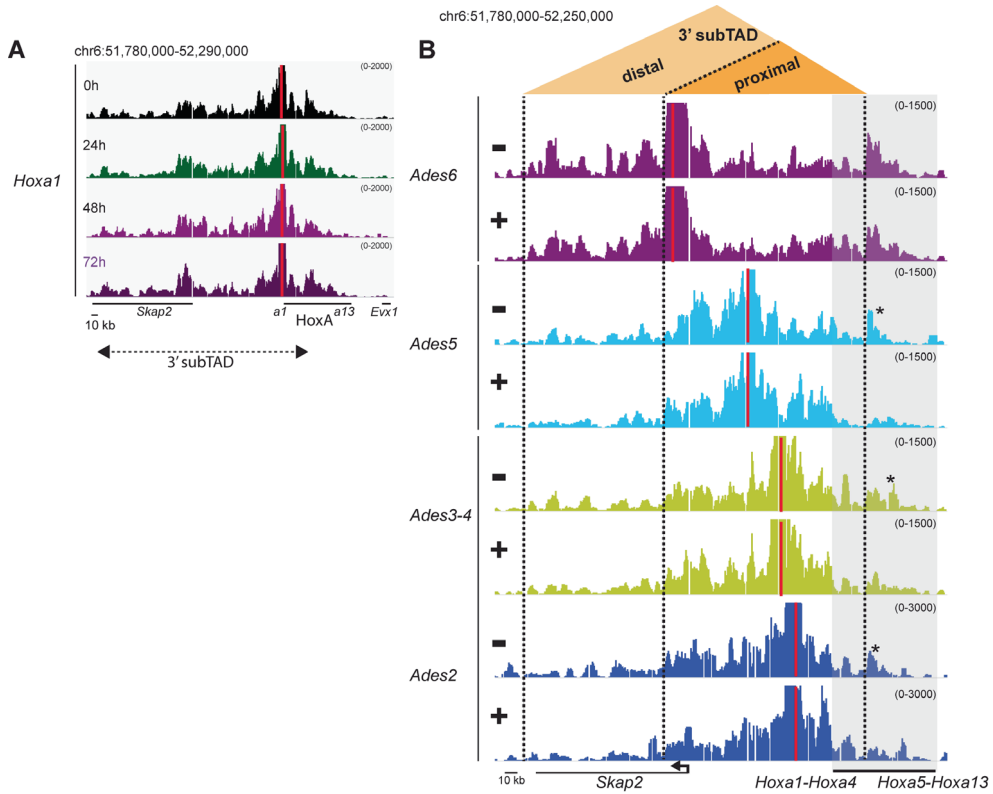
(Right page, bottom)

Supplemental Figure S4: Intensive contacts between a subregion of the Hox 3' TAD and 3' HoxA (A) Upper row, HiC profile-derived TAD positions (Dixon et al. 2012). Below, 4C-seq profiles of interactions viewed from different Ades enhancers and from *Hoxa1* and *Hoxa6*. (B) Zoom in on Ades enhancer interactions within the 3' subTAD. Red lines indicate 4C-seq viewpoints.

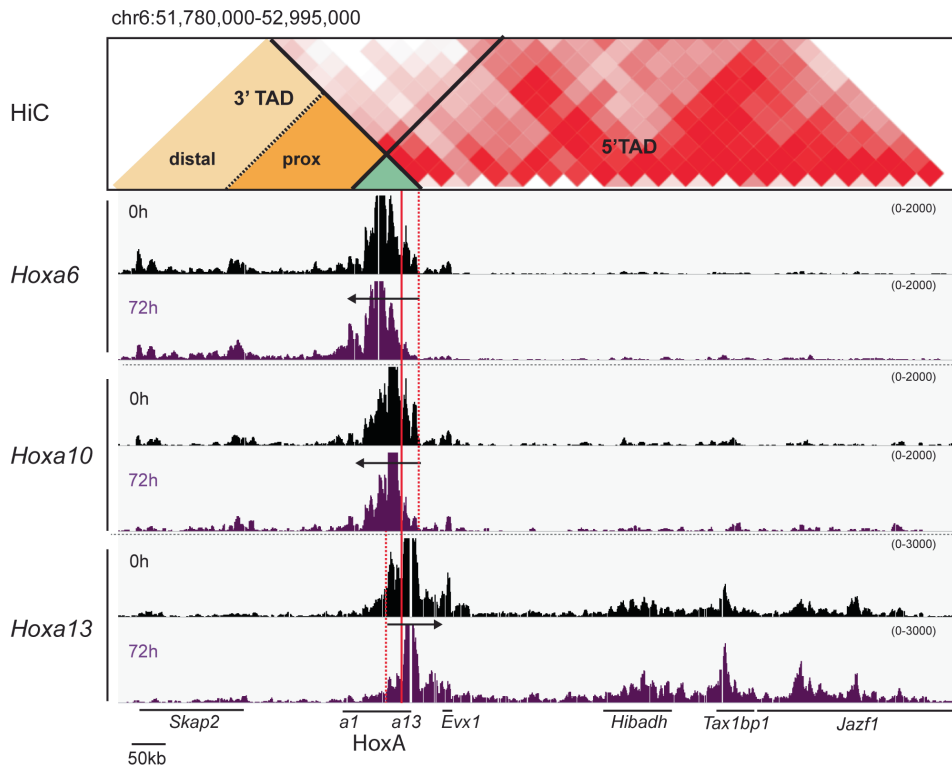


4

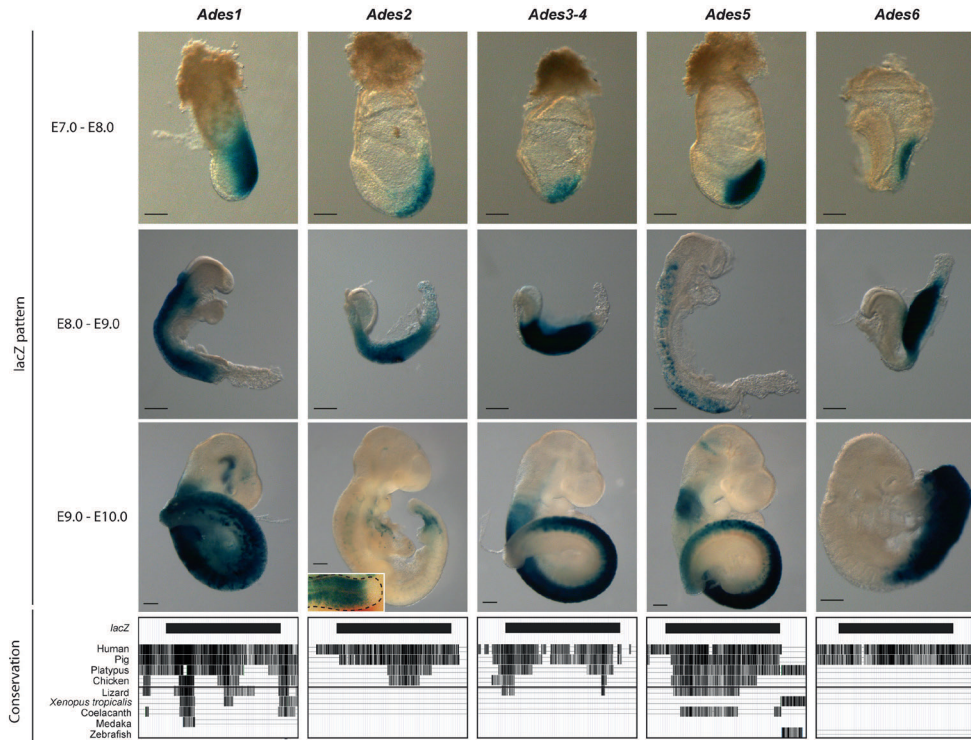




Supplemental Figure S5: Chromatin conformation dynamics in the 3' HoxA neighborhood during Wnt activation. (A) 4C-seq profiles of 3' Hox gene *Hoxa1* during activation. Chromosomal interactions of *Hoxa1* with the Ades region (between *Hoxa1* and *Skap2*) remain largely unchanged in both Hox transcriptionally inactive (0h) and Hox transcriptionally active situations (24 hours, 48 hours and 72 hours after Chiron stimulation). (B) 4C-seq profiles of Ades enhancers without (-) or with (+) Chiron administration. *Ades2*, *Ades3-4*, *Ades5* are located in the proximal part of 3' subTAD, and lose their contacts with the middle Hox genes (asterisks) upon 48h of Chiron activation. *Ades6* is located at an internal boundary in the 3' subTAD and its contact extends beyond the proximal subTAD domain. Upon Wnt activation (24h), contacts on the middle Hox genes and in the distal subTAD remain unchanged. Red lines indicate 4C-seq viewpoints.



Supplemental Figure S6: Chromatin conformation dynamics in the 5' HoxA neighborhood during Wnt stimulation. Top row, HiC profile-derived TAD positions (Dixon et al., 2012). 4C-seq profiles of middle Hox genes *Hoxa6* and *Hoxa10* and posterior *Hoxa13* over time of activation (72 hour Chiron). Both *Hoxa6* and *Hoxa10* further restrict their interactions to more 3' regions (away from the regions between red line and red dotted line). For *Hoxa13*, located in the 5' TAD, a decrease in frequency of contacts with the 3' TAD is observed (between red dotted line and red line).



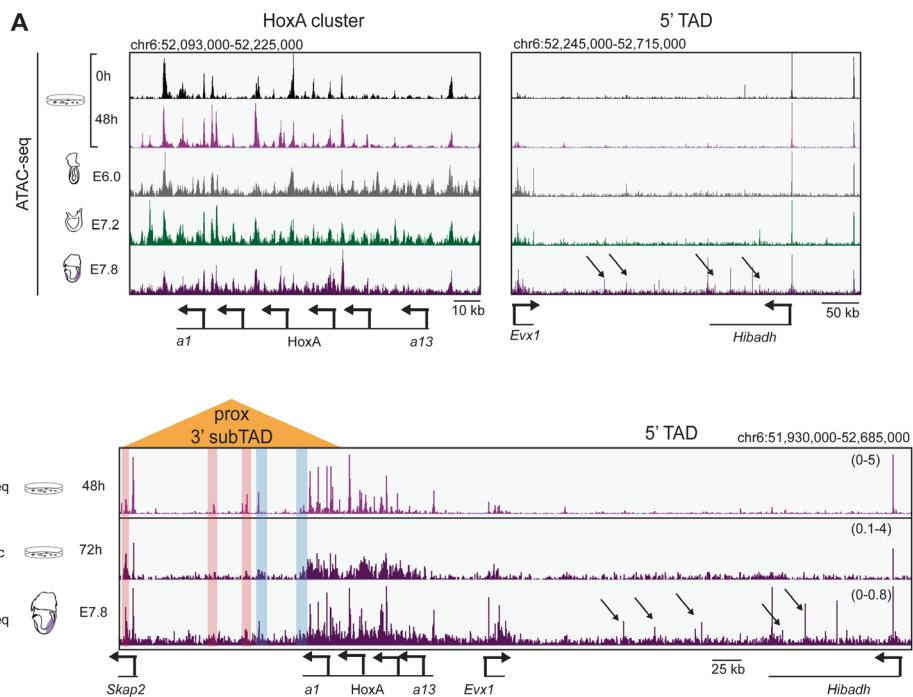
Supplemental Figure S7: Ades enhancer characteristics: domains of activity and evolutionary conservation. Up- per panel: *lacZ* domain of activity of Ades enhancers from E7.3 to mid-gestation. Insert in *Ades2*: E10.5 tailbud of another embryo showing posterior activity of *Ades2* at posterior levels. Lower panel: evolutionary conservation of Ades enhancers among vertebrates. The fragments used for the *lacZ* reporter assays are indicated by the black bars. Scale bars, 200 μm .

(Right page, on top)

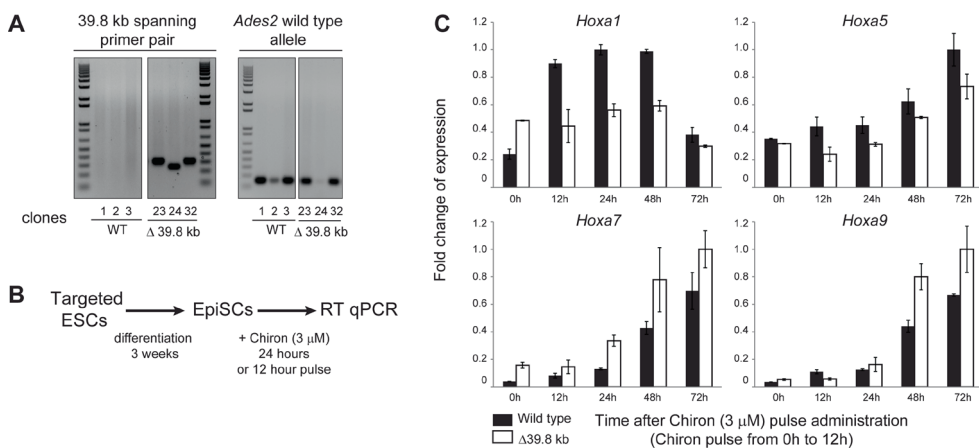
Supplemental Figure S8: DNA accessibility and enhancer acetylation throughout the *HoxA* landscape during Wnt-stimulated *HoxA* activation. (A) ATAC-seq profile along the *HoxA* cluster (left) and its 5' TAD (right panel). Arrows in the right panel point to the positions of the mid-gestation limb-specific enhancers (see text). (B) Comparison of ATAC-seq profile of Wnt-activated EpiSCs (48 hours Chiron), H3K27ac distribution in Wnt-activated EpiSCs (72 hours, *Wnt3* null EpiSCs), and ATAC-seq profile in posterior tissues of E7.8 embryos (lower lane, purple). Arrows as in A. Activation of regulatory elements after Wnt stimulation is entirely restricted to the 3' TAD (see Fig. 2A, 3B); DNA accessible positions at the 5' side of the cluster in embryos after *Hox* expression has started (E7.8), never become accessible or acetylated upon Wnt exposure in EpiSCs.

(Right page, bottom)

Supplemental Figure S9: CRISPR/Cas9-directed deletion of a proximal region containing *Ades1* and *Ades2* enhancers and its effect on *Hoxa1* and more 5' *Hoxa* genes. (A) Examples of PCR-based genotyping of wild type (clones 1, 2, 3) and alleles carrying the deletion of the Wnt-dependent regulatory region (clones 23, 24, 32). Clone 24 carries a homozygous deletion. (B) Differentiation protocol for obtaining ESCs-derived EpiSCs; wild type and targeted ESCs are induced to differentiate in parallel (see Material and methods). (C) Effects of the deletion on *Hoxa1*, *Hoxa5*, *Hoxa7* and *Hoxa9* over a 72 hour interval, after an initial 12-hour pulse of Chiron.



4



Supplemental Tables

All Supplemental Tables are available on genesdev.cshlp.org/content/30/17/1937/suppl/DC1 and are available from the author upon request.

Supplemental Table S1: Normalized Cq values of quantitative RT-PCR

Supplemental Table S2: Genotyping primers

Supplemental Table S3: Quantitative RT-PCR primers

Supplemental Table S4: 4C-seq primers

Supplemental Table S5: Ades regions cloning primers

Supplemental Table S6: Crispr/Cas9 sgRNA oligos

Supplemental Material and methods

Quantification of 4C-seq interactions

Normalized sequencing reads of the entire 3' TAD (chr6:51118074-52106126), and the proximal (chr6:51937875-52106126) and distal (chr6:51803027-51937875) part of the 3' subTAD were count. The *HoxA* cluster region is excluded from the quantification. Coordinates are from mm9 genome assembly.

Chiron pulse experiment on EpiSCs

Feeder-free ESC-derived EpiSCs (WT and containing the Δ 39.8 kb deletion) were cultured in presence of Wnt inhibitor IWP2. The medium was replaced by medium containing Chiron (3 μ M). After 12 hours, Chiron-containing medium was replaced by feeder-free medium without IWP2 to avoid continuous exogeneous Wnt activation.

Chapter 5

Cdx is a crucial player during colinear Hox activation and defines a trunk segment in the Hox cluster topology

Roel Neijts^{1*}, Shilu Amin^{1*}, Carina van Rooijen¹ and Jacqueline Deschamps¹

**Equal contribution*

¹Hubrecht Institute, Developmental Biology and Stem Cell Research, Uppsalalaan 8, 3584 CT Utrecht, and UMC Utrecht

Neijts R^{*}, Amin S^{*}, van Rooijen C and Deschamps J. Cdx is a crucial player during colinear Hox activation and defines a trunk segment in the Hox cluster topology,

In revision

**Equal contribution*

Abstract

Cdx and Hox transcription factors are important regulators of axial patterning and are required for tissue generation along the vertebrate body axis. Cdx genes have been demonstrated to act upstream of Hox genes in midgestation embryos. Here, we investigate the role of Cdx transcription factors in the gradual colinear activation of the Hox clusters. We found that Hox temporally colinear expression is severely affected in epiblast stem cells derived from Cdx null embryos. We demonstrate that after initiation of 3' Hox gene transcription, Cdx activity is crucial for H3K27ac deposition and for accessibility of *cis*-regulatory elements around the central – or ‘trunk’ – Hox genes. We thereby identify a Cdx-responsive segment of HoxA, immediately 5' to the recently defined regulatory domain orchestrating initial transcription of the first Hox gene. We propose that this partition of HoxA into a Wnt-driven 3' part and the newly found Cdx-dependent middle segment of the cluster, forms a structural fundament of Hox colinearity of expression. Subsequently to initial Wnt-induced activation of 3' Hox genes, Cdx transcription factors would act as crucial effectors for activating central Hox genes, until the last gene of the cluster arrests the process.

Introduction

Hox and Cdx genes, derived from an ancient ProtoHox gene cluster, are pivotal regulators of axial patterning and tissue generation along the vertebrate body axis (Krumlauf, 1994; Kmita and Duboule, 2003; Young and Deschamps, 2009). Mutations in Cdx genes lead to arrest of axial extension (Chawengsaksophak et al., 1997; Chawengsaksophak et al., 2004; van Rooijen et al., 2012) as a result of the exhaustion of axial progenitors in the posterior growth zone (Young et al., 2009; Neijts et al., 2014; Amin et al., in press). Cdx genes are key regulators of Hox genes, as their inactivation affects central Hox gene expression at midgestation, which in turn leads to homeotic transformations in the vertebral column (Subramanian et al., 1995; van den Akker et al., 2002). Interestingly, central – or ‘trunk’ – Hox genes *Hoxa5* and *Hoxb8* are able to rescue Cdx mutant truncation phenotypes (Young et al., 2009).

Recently we have used epiblast stem cells (EpiSCs), derived from the epiblast of pre-gastrulation embryos (embryonic day 6.0, E6.0), to study the molecular mechanism of Hox gene induction (Neijts et al., 2016). We have demonstrated that the HoxA cluster, and its primed early *cis*-regulatory landscape, is transcriptionally activated upon Wnt stimulation. Interestingly, after the initial activation of the earliest (3') Hox genes, a subsequent temporal colinear activation of the rest of cluster members takes place, down to the 5'-most Hox13 genes (Neijts et al., 2016). This temporal transcriptional dynamics is similar to the sequential activation observed in embryos between E7.2 and E9.0 [reviewed in (Deschamps and van Nes, 2005)]. Temporal Hox colinearity is widely present in the animal kingdom (Duboule, 2007). Different parameters have been proposed to contribute to bilaterian colinear Hox activation, including the intrinsic Hox locus organization (the cluster), the surrounding *cis*-regulatory landscape (enhancers, lncRNAs), active and inactive chromatin marks (histone modifications), boundary elements (like CTCF binding regions), and chromatin conformation (Noordermeer and Duboule, 2013).

We recently observed that, concurrently with 3' Hox activation, *Cdx2* becomes transcriptionally active in EpiSCs stimulated by Wnt signaling (Amin et al., in press). By performing ChIP-seq experiments on stimulated EpiSCs, we could identify regions bound by Cdx2 genome-wide. In agreement with their sensitivity to regulation by Cdx gene products, all four Hox clusters were found to be direct target of Cdx2, and trunk Hox gene expression was demonstrated by RNA-seq to be downregulated in mouse embryos lacking all three Cdx genes ('Cdx triple null') (Amin et al., in press).

Here, we investigate the role of Cdx in the temporal colinearity of Hox gene expression. As Cdx triple null embryos do not generate any post-occipital tissue as a result of failing to maintain their posterior growth zone and axial progenitor population (van Rooijen et al., 2012), it is difficult to study the Cdx-Hox interaction in the mutant embryo. Previously we have generated EpiSCs from Cdx null embryos (Amin et al., in press). In the present work we used these cells to study the transcription and chromatin dynamics of Hox loci during their Wnt-triggered activation in absence of Cdx. We now find that Cdx genes are required for temporal colinear Hox gene expression, and are crucial for the opening and activation of *cis*-

elements proximal to trunk Hox genes. Cdx gene products exert their function in a genomic segment that is different from two other distinguishable segments, the Wnt-dependent 3' part, and a 5' posterior part containing the late Hox13 gene.

Results

Cdx2 binds the Hox clusters and their regulatory landscapes in Wnt-stimulated EpiSCs

Cdx and 3' Hox genes are initially expressed in the posterior streak region of the E7.2 mouse embryo and follow the same spatiotemporal dynamics during axial elongation (Young and Deschamps, 2009). In EpiSCs, both 3' Hox genes and *Cdx2* start to be expressed upon Wnt stimulation (using the Wnt3a protein or the Wnt agonist CHIR99021 – hereafter 'Chiron') (Neijts et al., 2016; Amin et al., in press). We previously performed ChIP-seq experiments against Cdx2, the most dominant Cdx member during embryonic development, on Chiron/Fgf8-treated EpiSCs (24 hours of stimulation). We could identify almost 4000 Cdx2-bound regions, genome-wide. All four Hox clusters were found to be heavily occupied by Cdx2 (see Figure 1A)(Amin et al., in press).

Several of the sequences bound by Cdx2 have been demonstrated to be important during axial patterning, like elements close to *Hoxa5* (Tabaries et al., 2005), *Hoxa7* (Gaunt et al., 2004), *Hoxb8* (Charite et al., 1998) and *Hoxc8* (Shashikant and Ruddle, 1996) (Figure 1B). Specific deletion of the latter enhancer was shown to affect the timing of *Hoxc8* initial expression (Juan and Ruddle, 2003).

In addition to sequences within the four clusters, Cdx2 occupancy was also observed at elements that are part of the Hox *cis*-regulatory landscapes. Some of the early Hox enhancers at the 3' side of HoxA (*Ades1*, *Ades5* and *Ades6*) (Neijts et al., 2016) are bound by Cdx2 (Figure 1B). *LacZ* reporter assays demonstrated that these elements are all active in the posterior growth zone of the embryo, at least until E9.0-E9.5 (Neijts et al., 2016). In addition to these early enhancers, some late enhancers are also occupied by Cdx2. For instance, we find that the *e4* enhancer, 280 kb upstream of *Hoxa13* and active in limb and genitals (Berlivet et al., 2013), is bound by Cdx2 (data not shown). Moreover, the *cs38* element, 380 kb away from *Hoxd1* (Beccari et al., 2016) is occupied by Cdx2 as well (data not shown). This element is a transcription start site of *Hotdog* and *Twin of Hotdog* lncRNAs, important for emergence of the endoderm-derived caecum (Delpretti et al., 2013).

We next compared our Cdx2 binding data from induced EpiSCs with published ChIP-seq data obtained from Cdx2 overexpressing embryonic stem cells (ESCs)(Mahony et al., 2014). As expected, Cdx binding sites over the central Hox genes are present in both situations (Supplemental Figure S1). However, we observed that overexpressed Cdx2 in ESCs – and also in ESC-derived endoderm and motor neurons (Mahony et al., 2014) – binds proximal to late trunk Hox genes like *Hoxa7*, *Hoxa9*, *Hoxb8*, *Hoxc8* and *Hoxc10* relatively heav-

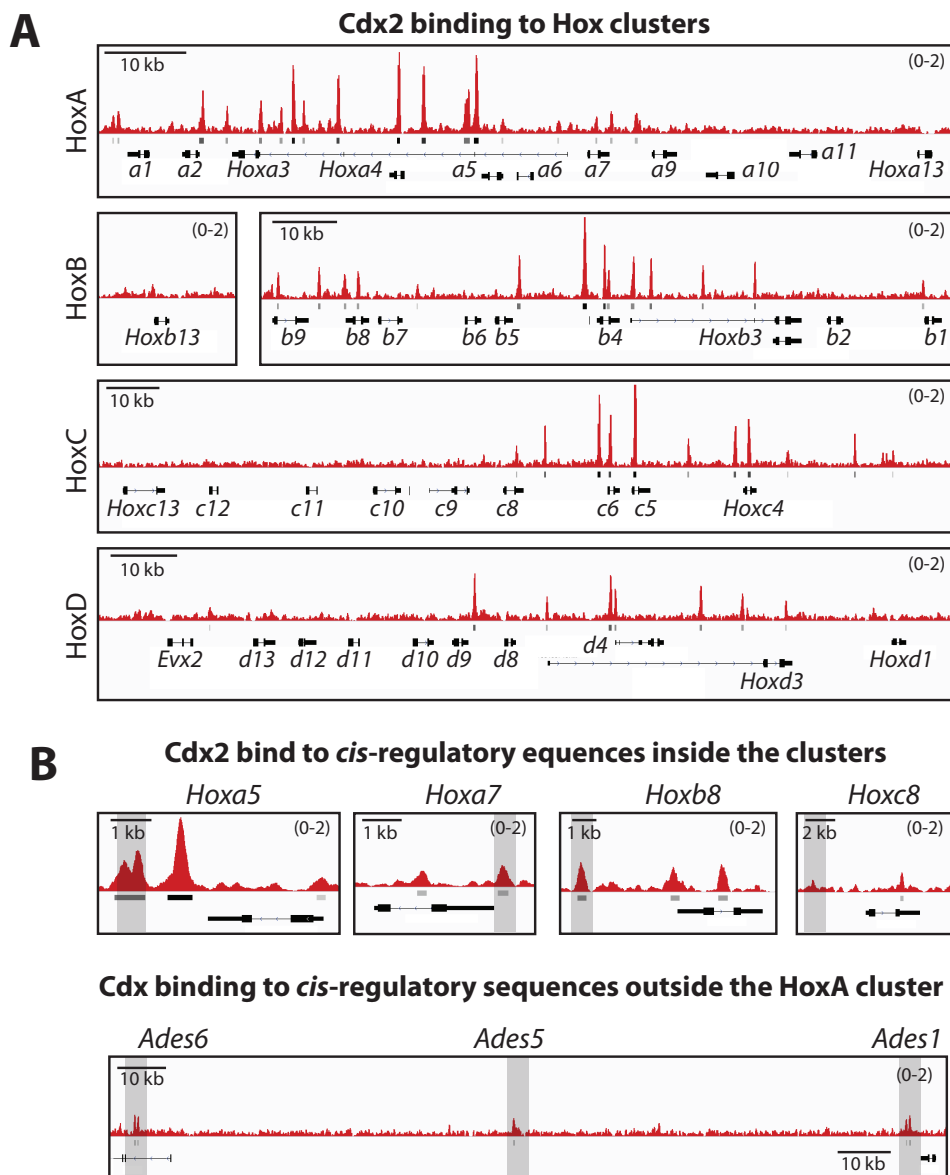


Figure 1: Distribution of Cdx2 binding throughout the Hox clusters in EpiSCs, induced for 24 hours with Chiron and Fgf8. A) Binding of Cdx2 on HoxA, HoxB, HoxC and HoxD clusters. B) Upper panels, Cdx2 binds to Cdx-dependent *cis*-regulatory elements in different Hox clusters (references in text). Lower panel, early HoxA enhancers *Ades1*, *Ades5* and *Ades6* are bound by Cdx2. The bars underneath the Cdx2 binding data represent the peaks that are identified by the MACS peak calling algorithm. (Data from Amin et al., in press.)

ily compared to the binding in EpiSCs (Supplemental Figure S1). On the contrary, induced EpiSCs show Cdx2 binding exclusively at the 3' part of the clusters which are not occupied in ESCs (Supplemental Figure S1, in orange bars).

A very contrasted situation therefore opposes the clear Cdx2 binding profile in the 3' part of HoxA in EpiSCs, to that in ESCs where Cdx2 does not bind. This is likely to be a consequence of the Wnt-driven activation of the 3' side of the cluster that would make Cdx2 binding possible. Strikingly, numerous Cdx2 binding sites are observed in the interval between *Hoxa3* and *Hoxa9* in both induced EpiSCs and Cdx2-overexpressing ESCs, with relatively more intense binding in the *Hoxa7-Hoxa9* interval in the ESC situation. A similar dichotomy is observed for the HoxB cluster (Supplemental Figure S1). This suggests that the middle part of the Hox cluster has a different status than the 3' part, which needs to have been sensitized by an initial Wnt activation (Neijts et al., 2016) before it accommodates Cdx2 binding.

Hox cluster colinearity of expression is impaired in Cdx mutant EpiSCs

Considering the functional evidence of Cdx regulatory activity on Hox gene expression during development, and Cdx2 binding on the clusters, we tested whether Cdx2 is involved in temporal colinear activation of Hox genes. We previously did not observe any defects in the onset of *Hoxb1* in Cdx triple null embryos at E7.5 (van Rooijen et al., 2012), whereas the more 5' located genes were not activated or were affected in their expression levels. However, this impaired expression was believed to be the result of the exhaustion of the posterior progenitor population in these mutant embryos occurring after E7.5.

Comparing the non-induced levels of HoxA transcripts in wild type (WT) and Cdx null EpiSCs revealed slightly higher levels of initial basal expression of Hox genes in mutant cells, and this feature is particularly true for *Hoxa5* (Figure 2A). Because both cell lines were cultured in the presence of IWP2 and in the absence of Wnt, the cause of this elevated basal expression in the Cdx null cells is elusive. We compared the inductibility of the HoxA genes upon Wnt pathway activation in Cdx triple null and WT EpiSCs. When we induced WT EpiSCs by Chiron over a time course of 72 hours, we observed a normal colinear activation of HoxA. The transcriptional levels of trunk HoxA genes did not gradually increase in Cdx null EpiSCs as in WT EpiSCs (Figure 2A) and the levels of *Hoxa5* and *Hoxa9* remained relatively constant. The 5'-most HoxA gene, *Hoxa13*, did not get activated as in WT conditions.

In uninduced Cdx triple null cells, *Hoxa5* exhibited a high level of the active histone mark H3K27ac, concurrent with the high initial transcriptional level (see arrow in Figure 2B, and Figure 2A). After Wnt stimulation, the levels of acetylation along the HoxA cluster were increased, and the activation domain was spread towards the 5' side in WT cells (Figure 2B). In the absence of Cdx, several positions that normally bind Cdx2 in the HoxA locus remain depleted of H3K27ac, in particular in the middle of the cluster. For the HoxB cluster the situation is even more striking (Figures 2C,D). Most of Wnt-induced acetylation in the Cdx null cells took place 3' to *Hoxb4*, whereas acetylation in the WT cells reached down to the *Hoxb9* locus, including the part of the cluster rich in Cdx2 binding (Figure 2D).

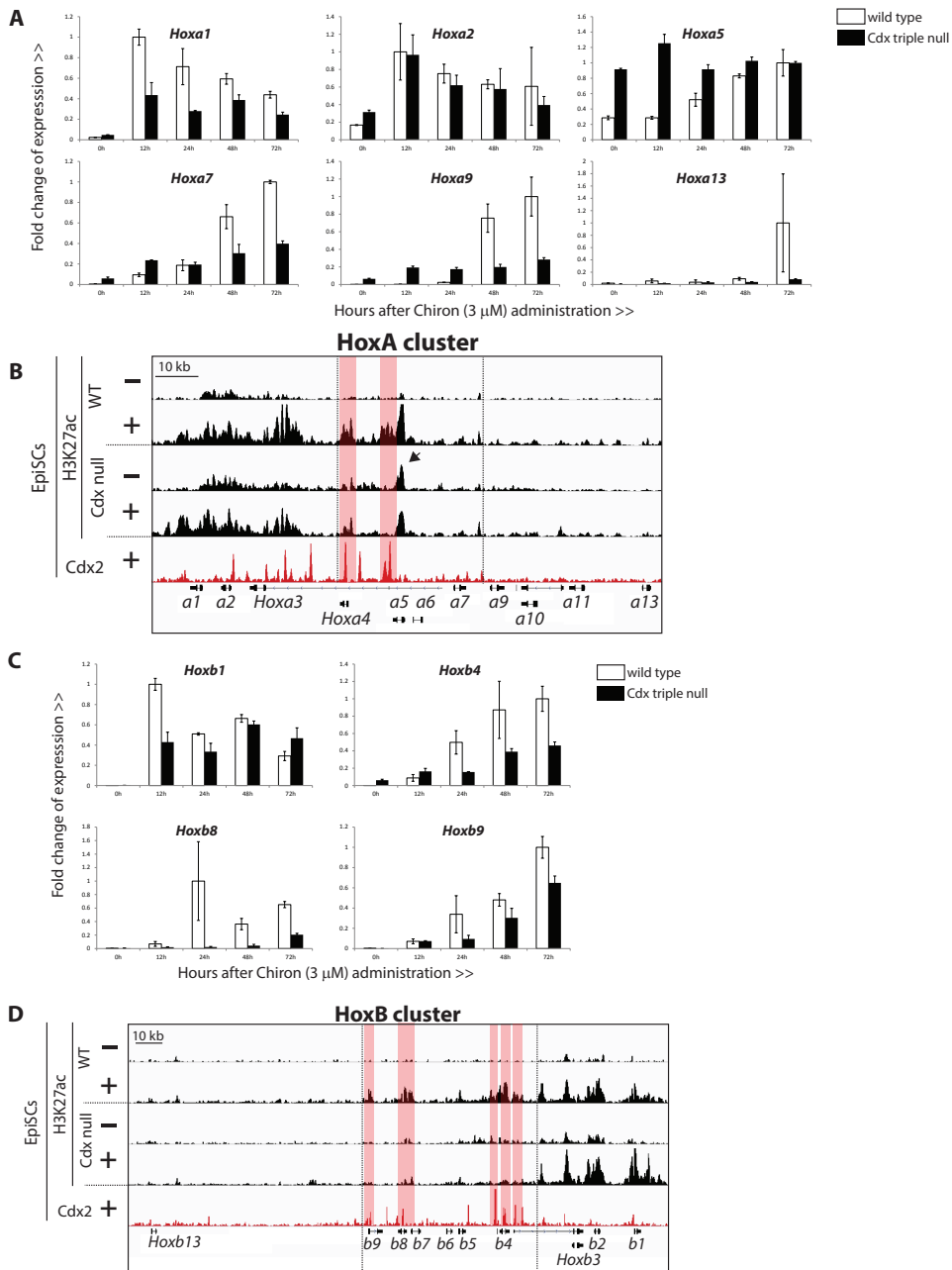


Figure 2: Temporal colinearity of Hox activation is affected in the absence of Cdx. A) HoxA genes are not colinearly upregulated in Cdx triple null EpiSCs (black), as they are in WT EpiSCs (white). Error bars indicate \pm S.D. B) An initial slight enrichment of H3K27ac is present at the HoxA cluster in Cdx mutant EpiSCs, compared to WT EpiSCs (in particular for *Hoxa5*, arrow). After 24 hours of stimulation, H3K27ac levels increase in WT cells and less in Cdx null cells (indicated by red bars). C) The regulation of HoxB genes is affected in Cdx triple null EpiSCs, compared to their regulation in WT EpiSCs. D) H3K27ac decoration in conditions as in B); here for the HoxB cluster.

Our data reveal that Cdx inactivation leads to disturbed colinearity of Hox gene induction and to a diminished H3K27ac deposition in the central part of the clusters, under Wnt-induced conditions.

Wnt-induced Hox activation opens Cdx2-binding positions in the middle part of the Hox clusters

To understand the mechanism by which Cdx2 influences colinear Hox gene activation, we examined the chromatin opening at Hox loci during Wnt stimulation. We performed an ‘assay for transposase-accessible chromatin’ (ATAC-seq) (Buenroostro et al., 2013) in non-induced WT EpiSCs. Multiple elements appear to already be accessible throughout the clusters, including in the gene body of Hox13 genes (see Figure 3A, for HoxA) (Neijts et al., 2016). The ATAC-seq profile over the cluster is very similar to the distribution of H3K4me3 – except for the 5’-most *Hoxa13* that lacks the active histone methylation mark. Apart from this *Hoxa13* exception, the deposition of active histone mark H3K4me3 by the trithorax group methyltransferases therefore correlates with the opening of the chromatin at these positions. The 5’ boundary of the H3K4me3 decoration – and of the opened chromatin domain – marks the boundary of the Cdx2 binding domain determined in induced EpiSCs (Figure 3A). The location of this virtual boundary is a well-conserved sequence (*CTCF Binding Site 5, CBS5*) shown to bind CTCF and Oct4 in ESCs and to play a role in Hox boundaries in motor neurons (Kim et al., 2011; Narendra et al., 2015)

Stimulation of EpiSCs by Wnt (Chiron for 48 hours), leads to further opening of the HoxA cluster and of the early enhancers *Ades1* and *Ades2* (Neijts et al., 2016). Comparison of the sequences opened upon this Wnt exposure with the regions bound by Cdx2 revealed as striking overlap (Figure 3B). In the entire *Ades1-CBS5* interval, sequences that become accessible upon induction appear to be bound by Cdx2 (green regions in Figure 3B). This suggests that Cdx2 immediately binds to newly opened chromatin, or alternatively, that Cdx2 acts as a pioneer transcription factor (Zaret and Mango, 2016) opening up these positions.

5

Chromatin opening around trunk Hox genes is Cdx-dependent

To functionally test whether Cdx is required for the opening of Hox *cis*-sequences, we performed ATAC-seq on Cdx null EpiSCs. ATAC-seq on non-induced Cdx mutant cells show the same distribution of opened chromatin positions throughout the four Hox clusters as in non-induced WT EpiSCs (data not shown).

Upon stimulating Cdx null EpiSCs for 48 hours with Chiron, we compared the chromatin opening profile with that in induced WT cells. Several positions within the Hox clusters are dependent on Cdx to be opened (Supplementary Figure S2A) [data from (Amin et al., in press)]. Interestingly, we observed a dichotomy in the opening of Hox regions. In Cdx null EpiSCs, the 3’ part of the HoxA cluster is open similarly and at the same positions as in WT cells (Figure 4, orange bars). This is true even for sequences that are bound by Cdx2. The 3’ part of the Hox cluster is thus not dependent on Cdx for becoming open. In contrast, ele-

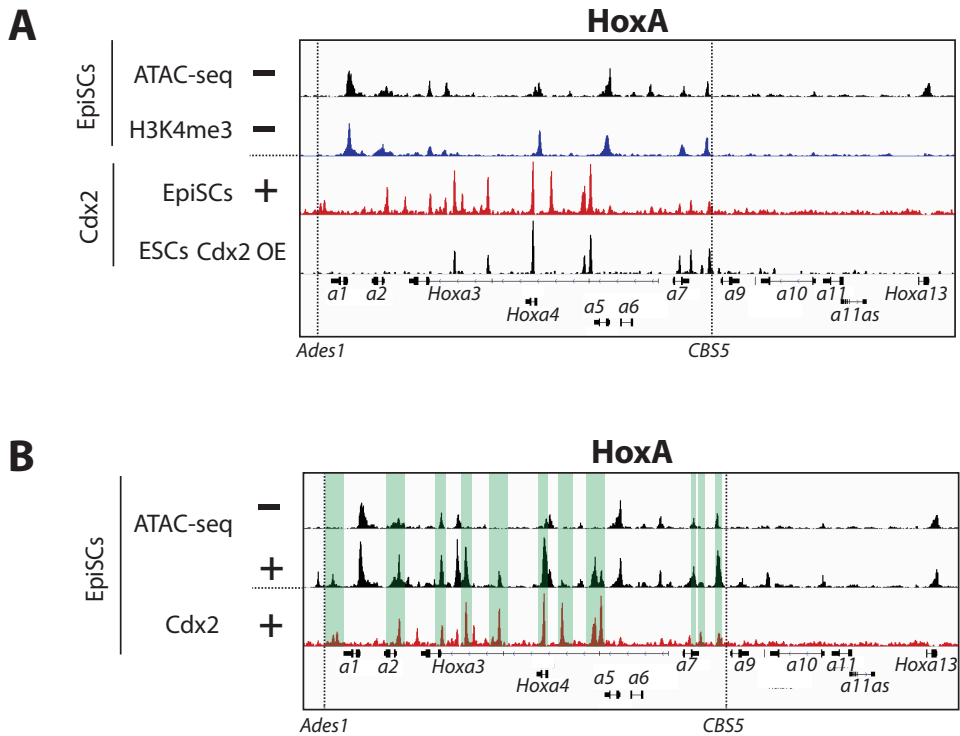


Figure 3: DNA accessibility and Cdx2 distribution before and during Hox activation in EpiSCs. A) The DNA accessibility (ATAC-seq) profile (black) is similar to the distribution of H3K4me3 (blue), except for *Hoxa13*, in non-induced EpiSCs. The distribution of Cdx2 binding (red for induced EpiSCs and black for Cdx2-overexpressing ESCs) and H3K4me3 decoration, together with the DNA accessibility, share the same 5' boundary (dashed line) at *CTCF binding site 5 (CBS5)* (Kim et al., 2011). B) During Wnt activation (48 hours, Chiron), DNA accessibility increases at the *HoxA* cluster. The positions that become accessible are the same that are bound by Cdx2 (in red) in induced EpiSCs (regions indicated by green bars).

ments in the middle part of the clusters are dependent on the presence of Cdx to become accessible. All positions that are bound by Cdx2 in the *Hoxa3-CBS5* interval in the WT cells have reduced levels of chromatin opening in Cdx mutant EpiSCs (Figure 4, purple bars). Such a Cdx-dependence for chromatin opening was previously seen for several non-Hox Cdx2-bound enhancers (Amin et al., in press). As expected, H3K27ac data show that some of the Cdx-dependent regions are affected in their acetylation levels in Cdx mutant cells (arrow in Figure 4, lower lanes). Noteworthy, all elements that are bound by Cdx2 in ESCs (Mahony et al., 2014) within the trunk *HoxA* area, seem to be dependent on Cdx for activation (Figure 4). It is not ruled out that Cdx2 is able to bind to more 5' *HoxA* genes in the *Hoxa9-Hoxa13* interval in other conditions of induction of EpiSCs. A more complete *Hox* cluster opening might be required for more 5'-located Cdx2 binding.

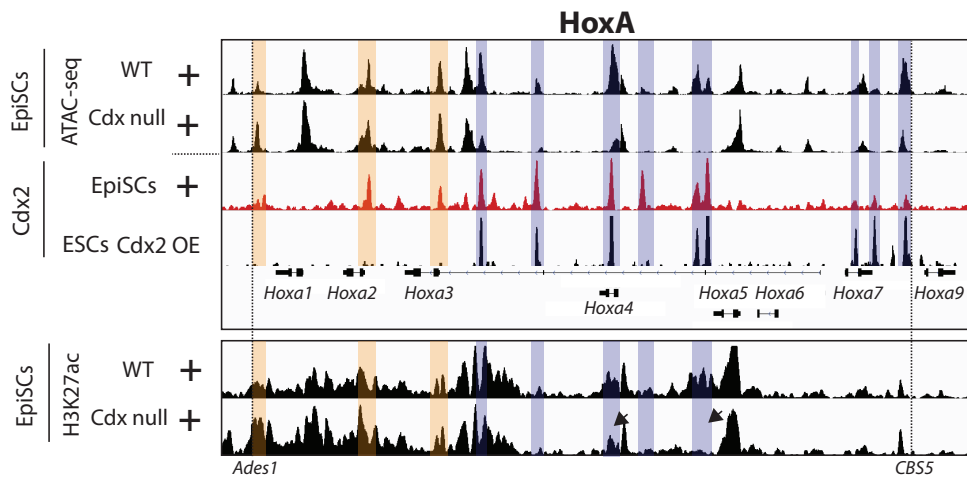


Figure 4: Cdx is required for the DNA accessibility of regulatory elements around trunk Hox genes. Upper panel: Stimulation of WT EpiSCs leads to opening of genomic positions at both 3' Hox genes and trunk Hox genes (see also Figure 3). In Cdx triple null EpiSCs, the 3' positions are open (orange bars) whereas elements around trunk Hox genes do not become accessible (purple bars). These Cdx-dependent positions overlap with regions that bind Cdx2 in Cdx-overexpressing ESCs (lower lane, in black). Lower panel: H3K27ac profiles of WT and Cdx null EpiSCs at HoxA show decreased levels of acetylation in Cdx mutant cells, at the positions that are dependent on Cdx for DNA accessibility.

For the HoxB cluster we find similar striking effects of Cdx-dependence. Our data reveals a need of Cdx for DNA opening and H3K27ac deposition for all genes located 5' of *Hoxb3* (arrows and purple bars in Supplemental Figure S2B).

Together, the ATAC-seq data indicate that Cdx transcription factors are required for trunk Hox gene opening and activation, and suggest that the Cdx gene products acts as pioneer transcription factors at these positions. The 3' part of the cluster is independent of Cdx for DNA accessibility.

Segmental distribution of Cdx2 binding and of Cdx-dependent chromatin opening in the HoxA cluster

Examination of the higher order architecture of the HoxA cluster, by chromosome conformation capture techniques as 4C-seq and HiC revealed that the locus is located at the boundary of two topologically associating domain (TADs) (Dixon et al., 2012). In superposition of global dichotomous organization, the HoxA locus is segmented in multiple subdomains (Figure 5A). The 3' part of the cluster (*Hoxa1-Hoxa4*) is located in a 3' subTAD, which includes the early enhancers; the central Hox genes (*Hoxa5-Hoxa10/11*) form a second segment. *Hoxa13* and its distant 5' enhancers form a third HoxA domain and are mainly in contact with the

large 5' TAD (Neijts et al., 2016), which is involved in late Hox regulation in limb and genitals (Berlivet et al., 2013; Lonfat et al., 2014) (not shown in Figure 5A).

Interestingly, we observed that both trunk (like *Hoxa6*) and posterior Hox genes that are poor in contacts with the 3' subTAD, could reach a very distant (1 mega base remote) region at the 3' boundary of the 3' TAD (hereafter: 'TAD boundary element', *TBE*) (Figure 5A). A 4C-seq view from the *TBE* reveals that it indeed interacts with the trunk and posterior Hox segment, and less with the 3' Hox genes and their 3' subTAD (Figure 5A). Interestingly, the activity domain of *TBE* in the embryo is very similar to that of trunk Hox genes, as it is initiated relatively late in the posterior streak and remains active in the posterior growth zone (Figure 5B). Notably, the *TBE* is not bound by *Cdx2*.

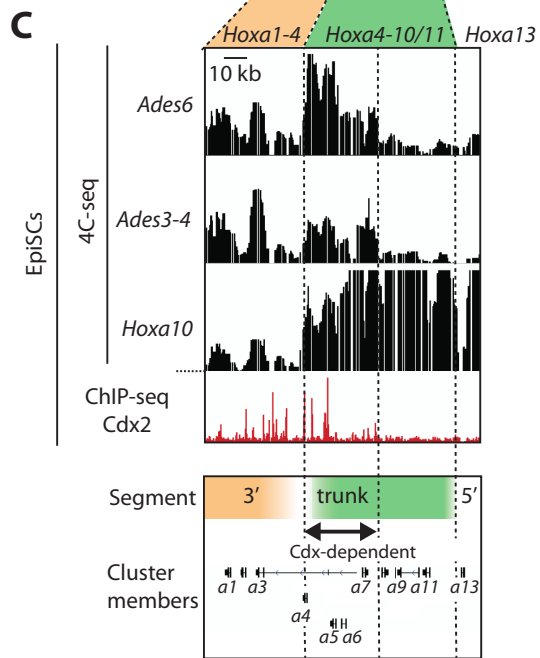
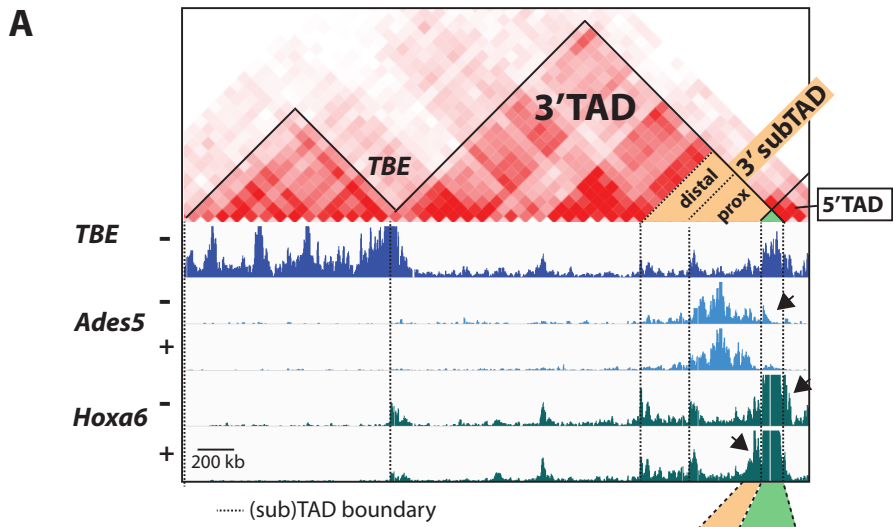
We wondered whether the distribution of the *cis*-elements – the opening of which is dependent on *Cdx* – is organized in a similar segmented manner as the locus. Therefore, we examined the interactions seen from 4C-seq viewpoints located in the same TAD in which the *Cdx2* binding sites are located. To avoid a too high number of interactions due to proximity, we selected three viewpoints relatively distant from the trunk Hox sequences. At the 5' side, we selected *Hoxa10* – which lies at the 5' limit of the 3' TAD. At the 3' side we selected the *Ades3-4* and *Ades6* enhancers – which are within the 3' subTAD, and at the 5' boundary of the 3'subTAD, respectively (Neijts et al., 2016).

Seen from *Ades6* and *Ades3-4*, a dense interaction domain overlaps with the *Cdx2* binding domain between *Hoxa4* and *CBS5* in uninduced conditions (Figure 5C). Not much interaction was scored 5' of *CBS5*, indicating that boundary might be functional in EpiSCs as it is in ESCs (Kim et al., 2011). From the *Hoxa10* viewpoint, we observed the same sharp boundary around *Hoxa4*. At the 3' side of this boundary, a relatively low amount of contacts are observed along a short 'low intensity' domain. Further 3', a higher intensity of contacts is observed, also seen from the other viewpoints. The latter domain includes the 3' part of the HoxA cluster, in which *cis*-elements do not require *Cdx* to be opened. 5' of the *Hoxa4* boundary, extremely high interactions with the *Hoxa10* viewpoint are observed, throughout the entire *Hoxa4-Hoxa13* interval (Figure 5C) – as observed from the *TBE* viewpoint (Figure 5A).

Collectively, we can distinguish differentially *Cdx*-bound domains in HoxA that overlap with the topological segments. At the 3'-most end of the cluster, a *Cdx*-independent domain is located spanning *Hoxa1* to *Hoxa3* (Figure 4B). Between *Hoxa4* and *Hoxa9* a central domain is identified, with sharp topological boundaries and with *cis*-elements that are fully dependent on *Cdx* for opening (see Figure 4). The most 5' part of the HoxA cluster (*Hoxa10-Hoxa13*) is activated much later than the 3' and central Hox genes. Its dependence on *Cdx*, shown in *Cdx* null EpiSCs, is probably secondary to the activation of central Hox genes, as no pre-existing binding sites are detectable at 24 hours of induction in EpiSCs.

Discussion

The progressive expression in time of the genes of the Hox clusters is of essential importance for the correct generation and antero-posterior patterning of the embryonic trunk in verte-



5

brates. We previously deciphered the mechanism of initial activation of the anterior (3') part of the HoxA cluster by early embryonic Wnt signals. We therefore made use of epiblast stem cells (EpiSCs) which, when induced by Wnt, are a good model for the posterior part of the early embryo where Hox genes are activated. Comparing the epigenetic and architectural features, and the expression of the clustered Hox genes in wild type and Cdx null EpiSCs, we could now study the interplay between Cdx and the Hox genes during transcriptional activation of the clustered Hox genes.

In this work we show that Cdx gene products intervene to shape the epigenetic landscape of the middle part of the HoxA cluster and to activate central Hox genes after the 3' genes have been turned on. This stepwise activation of the clustered Hox genes appears as at least one of the molecular genetic fundamentals of the temporal colinearity of Hox gene expression. Early embryonic signaling by Wnt3 in the posterior embryonic epiblast induces the expression of 3' Hox and of Cdx genes. Cdx gene products then activate the expression of the central and more 5' located Hox genes. The genomic area of action of the Wnt signal is the *cis*-regulatory landscape forming a 3' Hox subTAD (Neijts et al., 2016), whereas the domain activated by Cdx proteins is the central segment of the cluster, 5' to *Hoxa4*, where many Cdx2 binding sites have been detected. These successive phases of the Hox gene regulation are at least partially independently driven, as transcriptional initiation of *Hoxa1* is independent of Cdx, and the central HoxA genes need Cdx to be activated.

The sequentially of expression of the 3' and central Hox genes is crucial for the correct differential morphogenetic activity the Hox genes. The timing of initial transcription of *Hoxa1* and of the central Hox genes in the posterior-most epiblast is a determinant of the subsequent expression of the genes in the anterior part of the primitive streak where progenitors for the axis are located (Wilson et al., 2009). A gene initially transcribed early will be expressed early in the axial progenitor region, and thus will instruct the emerging axial structures at a more anterior level than in the case of a gene initially transcribed later (Forlani et al., 2003; Deschamps and van Nes, 2005).

The biological activity of Cdx gene products on trunk axial tissues is twofold. Central Hox genes and their Cdx activators ensure posterior growth by maintaining axial progenitors active in the growth zone. Partial or complete inactivation of the Cdx genes causes arrest of axial extension at levels posterior to the occipital part of the axis, and central Hox genes can rescue this effect (Young et al., 2009; van Rooijen et al., 2012). Cdx genes are also modulating

Figure 5: The HoxA locus can virtually partitioned in several segments, differentially bound by Cdx2. A) Overview of the HoxA segments: interactions between HoxA and its 3' surrounding seen from several viewpoints, with (+) or without (-) Wnt stimulation, and a distant element TBE. The *Hoxa6* and *Ades5* 4C-seq data is from Neijts et al., 2016. The HiC data is from promoter.bx.psu.edu. B) *LacZ* staining of embryos carrying the *TBE-lacZ* transgene. A, anterior. P, posterior. C) Distribution of Cdx2 binding (in red) in the different topological HoxA segments (divided by dashed lines). Elements in the 3' subTAD (in yellow) are not dependent of Cdx for their opening, whereas elements in the next topological segment (in green) depend on Cdx to be accessible.

antero-posterior identity of the tissues derived from the trunk progenitors at least in part by regulating central Hox genes (Charite et al., 1998; van den Akker et al., 2002; Tabaries et al., 2005). Some at least of the *cis*-elements activated by Cdx factors within the middle segment of the HoxA cluster correspond to identified transcriptional enhancers, as it is the case for a Cdx-responsive *Hoxa5* element and for *Hoxb8* upstream sequences (Charite et al., 1998; Tabaries et al., 2005). These elements, as we now show, require Cdx proteins for their accessibility and activation, and thus constitute the controllers of the central Hox gene activation necessary for regulating posterior growth and patterning of the trunk.

One particular of these central Hox gene controllers is the inter-TAD boundary element *TBE*, marking the 3' end of the 3' HoxA TAD. This element which lies 1 mega base 3' of the HoxA cluster does specifically contact the central HoxA genes, as revealed in 4C-seq experiments, and it does exhibit enhancer activity reproducing the expression of trunk HoxA genes in *lacZ* transgenic embryos *in vivo*. As it is accessible already in uninduced EpiSCs and was not seen to bind Cdx2 in our experimental conditions (data not shown), it may belong to an additional regulatory circuit controlling central HoxA genes from the remote inter-TAD position.

The genomic domain corresponding to the central Hox genes where Cdx2 was shown to bind and where it modifies the epigenetic status, extends between *Hoxa4* and *Hoxa9* in our experiments with Wnt-induced EpiSCs. We cannot rule out that genomic positions in HoxA posterior to *Hoxa9* also would bind Cdx2 in conditions of longer exposure to Wnt or other posterior signals in embryos *in vivo*. These positions would activate more 5' HoxA genes to sustain growth and patterning of more caudal trunk tissues. The *Hoxa13* gene, in any case, plays a distinct role, as Hox13 proteins antagonize Cdx2 by binding to the same targets, and so doing arrests axial elongation (Young et al., 2009; Amin et al., in press). The *Hoxa13* gene and the Hoxa13 protein therefore constitute the most posterior module ending up central Hox gene temporally colinear activation, possibly by direct competition with the trunk effector Cdx. This event also has to be timely regulated to operate the trunk to tail transition at the correct axial level.

We present a model of the gradual regulation of the different architectural units by their corresponding *trans*-acting effectors, which together direct the temporally colinear turning on of HoxA genes until Hoxa13 arrests the process (Figure 6).

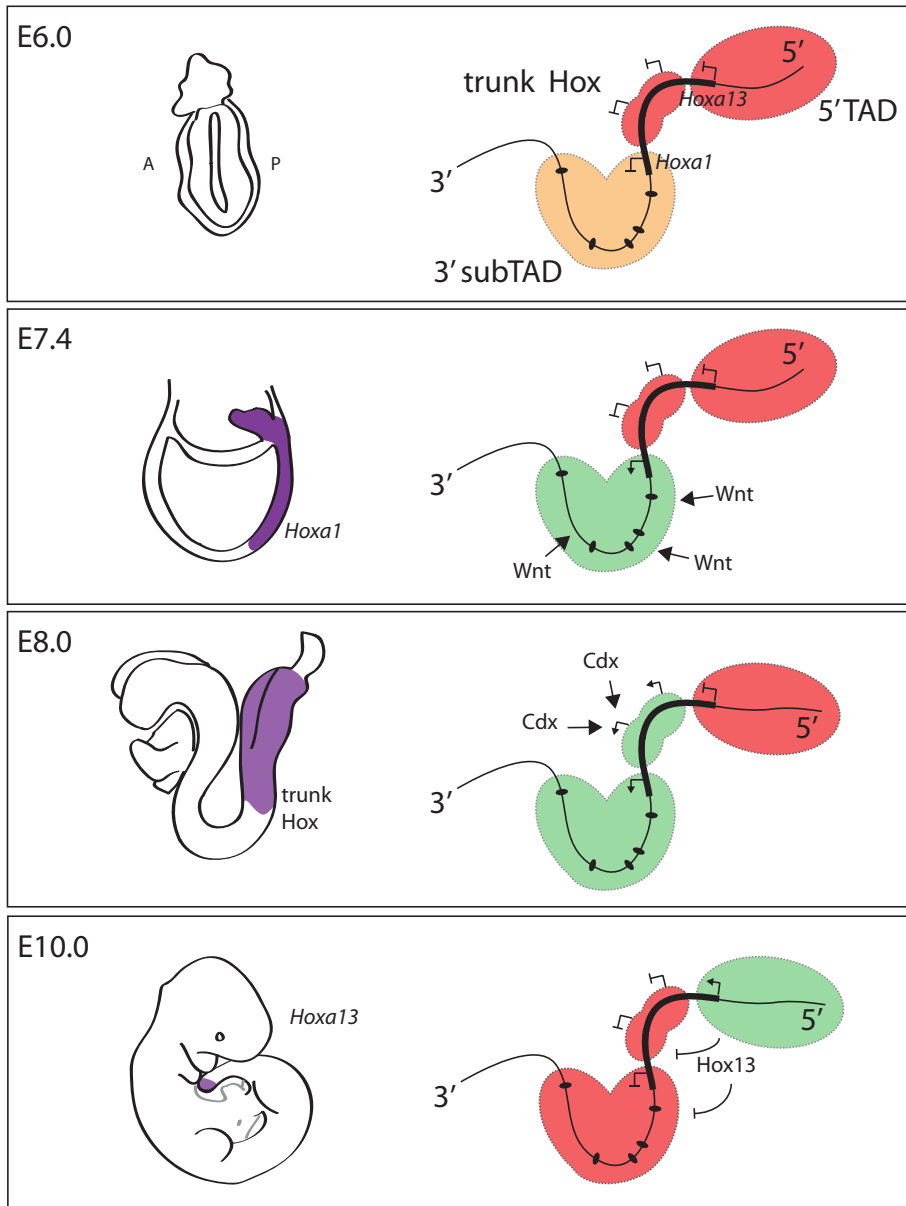


Figure 6: The gradual activation of HoxA, segment by segment, is progressively regulated by Wnt and Cdx. In E6.0 embryos no Hox genes are expressed; the 3' subTAD including the 3' regulatory domain (with enhancers, black dots) and *Hoxa1* are primed to be active (in orange). After embryos express Wnt posteriorly, *Hoxa1* is induced (at E7.2) and its expression is in the primitive streak region and allantois (in purple) at E7.4. The 3' subTAD is responsive to Wnt signals and is active. At E8.0 the trunk Hox genes are expressed. Trunk Hox genes correspond to the central topological domain (green in Figure 5), of which the first part was shown to be induced by Cdx in our experiments. Later on (E10.0), the 5'-most Hox segment is activated, and *Hoxa13* antagonizes the inducing effect of Cdx on trunk Hox genes. Purple, Hox gene expression domain. Orange, primed for activity. Red, inactive. Green, active.

Material and methods

EpiSCs

The generation and culture of wild type and Cdx triple null EpiSCs, and Wnt activation by Chiron, are described elsewhere (Neijts et al., 2016; Amin et al., in press).

Data sets

Data sets from ChIP-seq, ATAC-seq and 4C-seq experiments are from Gene Expression Omnibus series GSE81203 (Neijts et al., 2016), GSE84899 (Amin et al., in press) and GSE39435 (Mahony et al., 2014). HiC data is obtained from promoter.bx.psu.edu (Dixon et al., 2012). Experimental procedures are described in above mentioned references.

LacZ reporter assay

To generate a *TBE-lacZ* construct, a 6.2 kb fragment was cloned upstream of a minimal *Hsp68* promoter (Pennacchio et al., 2006). Forward primer: CACCGAGGTCCAGAAACGGGATTT. Reverse primer: AGAATTTCGCCATCAGGAGAC. Micro-injection and X-gal staining were performed as described elsewhere (Neijts et al., 2016).

RT qPCR

RNA isolation, cDNA synthesis and RT qPCR methods are describes elsewhere (Neijts et al., 2016).

4C-seq

The 4C-seq procedure was described elsewhere (Neijts et al., 2016). For the *TBE* viewpoint, the following primers were used. Forward primer: AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTGGTGACTGGAACCGTGAT. Reverse primer: CAAGCAGAAGACGGCATAACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTCATAAAAGGGA ACTATGCGT.

ChIP-seq

H3K27ac ChIP-seq was performed on non-induced and induced (24 hours Chiron) WT and Cdx null EpiSCs as described elsewhere (Neijts et al., 2016).

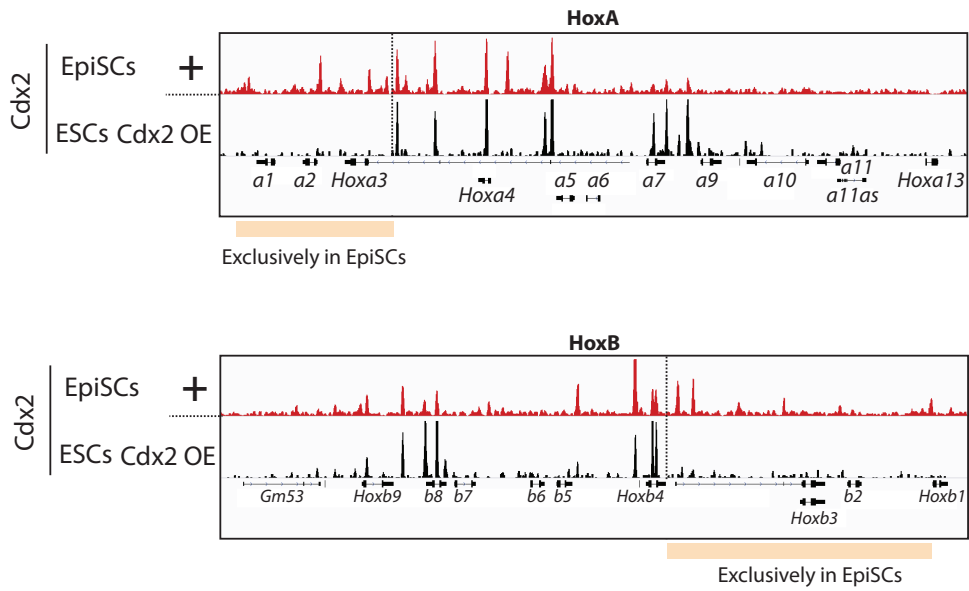
Acknowledgments

The authors Sander Tan and Geert Geeven for bioinformatical support on the H3K27ac ChIP-seq and 4C-seq data, respectively. We thank Menno P. Creyghton and Wouter de Laat for their expert advice on the ChIP-seq and 4C-seq approach, respectively. We thank Jeroen Korving for micro-injection of the *TBE-lacZ* construct. This work was supported by the Netherlands Institute for Regenerative Medicine (NIRM; grant FES0908).

References

- Amin S, Neijts R, Simmini S, van Rooijen C, Tan S, Kester L, van Oudenaarden A, Creighton MP, Deschamps J. in press. Cdx and T Brachyury co-activate growth signaling in the embryonic axial progenitor niche
- Beccari L, Yakushiji-Kaminatsui N, Woltering JM, Necsulea A, Lonfat N, Rodriguez-Carballo E, Mascrez B, Yamamoto S, Kuroiwa A, Duboule D. 2016. A role for HOX13 proteins in the regulatory switch between TADs at the HoxD locus. *Genes Dev* 30:1172-1186.
- Berlivet S, Paquette D, Dumouchel A, Langlais D, Dostie J, Kmita M. 2013. Clustering of tissue-specific sub-TADs accompanies the regulation of HoxA genes in developing limbs. *PLoS Genet* 9:e1004018.
- Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* 10:1213-1218.
- Charite J, de Graaff W, Consten D, Reijnen MJ, 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.
- Chawengsaksophak K, de Graaff W, Rossant J, Deschamps J, Beck F. 2004. Cdx2 is essential for axial elongation in mouse development. *Proc Natl Acad Sci U S A* 101:7641-7645.
- Chawengsaksophak K, James R, Hammond VE, Kontgen F, Beck F. 1997. Homeosis and intestinal tumours in Cdx2 mutant mice. *Nature* 386:84-87.
- Delpretti S, Montavon T, Leleu M, Joye E, Tzika A, Milinkovitch M, Duboule D. 2013. Multiple enhancers regulate Hoxd genes and the Hotdog LncRNA during cecum budding. *Cell Rep* 5:137-150.
- Deschamps J, van Nes J. 2005. Developmental regulation of the Hox genes during axial morphogenesis in the mouse. *Development* 132:2931-2942.
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485:376-380.
- Duboule D. 2007. The rise and fall of Hox gene clusters. *Development* 134:2549-2560.
- Forlani S, Lawson KA, Deschamps J. 2003. Acquisition of Hox codes during gastrulation and axial elongation in the mouse embryo. *Development* 130:3807-3819.
- Gaunt SJ, Cockley A, Drage D. 2004. Additional enhancer copies, with intact cdx binding sites, anteriorize Hoxa-7/lacZ expression in mouse embryos: evidence in keeping with an instructional cdx gradient. *Int J Dev Biol* 48:613-622.
- Juan AH, Ruddle FH. 2003. Enhancer timing of Hox gene expression: deletion of the endogenous Hoxc8 early enhancer. *Development* 130:4823-4834.
- Kim YJ, Cecchini KR, Kim TH. 2011. Conserved, developmentally regulated mechanism couples chromosomal looping and heterochromatin barrier activity at the homeobox gene A locus. *Proc Natl Acad Sci U S A* 108:7391-7396.
- Kmita M, Duboule D. 2003. Organizing axes in time and space; 25 years of colinear tinkering. *Science* 301:331-333.
- Krumlauf R. 1994. Hox genes in vertebrate development. *Cell* 78:191-201.
- Lonfat N, Montavon T, Darbellay F, Gitto S, Duboule D. 2014. Convergent evolution of complex regulatory landscapes and pleiotropy at Hox loci. *Science* 346:1004-1006.
- Mahony S, Edwards MD, Mazzoni EO, Sherwood RI, Kakumanu A, Morrison CA, Wichterle H, Gifford DK. 2014. An integrated model of multiple-condition ChIP-Seq data reveals predeterminants of Cdx2 binding. *PLoS Comput Biol* 10:e1003501.
- Narendra V, Rocha PP, An D, Raviram R, Skok JA, Mazzoni EO, Reinberg D. 2015. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science* 347:1017-1021.
- Neijts R, Amin S, van Rooijen C, Tan S, Creighton MP, de Laat W, Deschamps J. 2016. Polarized regulatory landscape and Wnt responsiveness underlie Hox activation in embryos. *Genes Dev* 30:1937-1942.
- Neijts R, Simmini S, Giuliani F, van Rooijen C, Deschamps J. 2014. Region-specific regulation of posterior axial elongation during vertebrate embryogenesis. *Dev Dyn* 243:88-98.

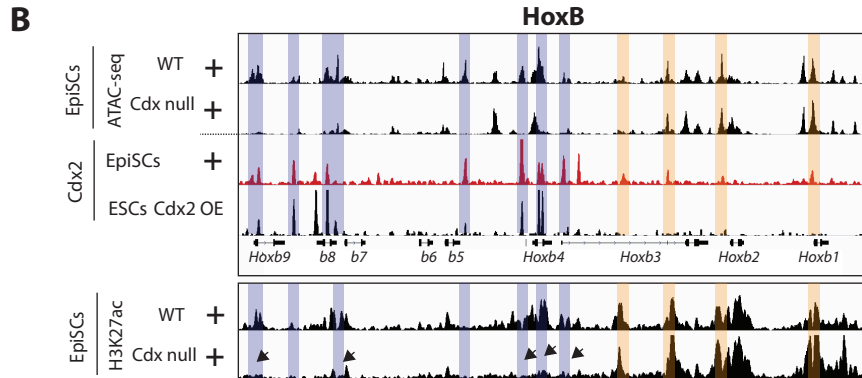
- Noordermeer D, Duboule D. 2013. Chromatin architectures and Hox gene collinearity. *Curr Top Dev Biol* 104:113-148.
- Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, Shoukry M, Minovitsky S, Dubchak I, Holt A, Lewis KD, Plajzer-Frick I, Akiyama J, De Val S, Afzal V, Black BL, Couronne O, Eisen MB, Visel A, Rubin EM. 2006. In vivo enhancer analysis of human conserved non-coding sequences. *Nature* 444:499-502.
- Shashikant CS, Ruddle FH. 1996. Combinations of closely situated cis-acting elements determine tissue-specific patterns and anterior extent of early Hoxc8 expression. *Proc Natl Acad Sci U S A* 93:12364-12369.
- Subramanian V, Meyer BI, 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.
- Tabaries S, Lapointe J, Besch T, Carter M, Woollard J, Tuggle CK, Jeannotte L. 2005. Cdx protein interaction with Hoxa5 regulatory sequences contributes to Hoxa5 regional expression along the axial skeleton. *Mol Cell Biol* 25:1389-1401.
- van den Akker E, Forlani S, Chawengsaksophak K, de Graaff W, Beck F, Meyer BI, Deschamps J. 2002. Cdx1 and Cdx2 have overlapping functions in anteroposterior patterning and posterior axis elongation. *Development* 129:2181-2193.
- 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.
- Wilson V, Olivera-Martinez I, Storey KG. 2009. Stem cells, signals and vertebrate body axis extension. *Development* 136:1591-1604.
- Young T, Deschamps J. 2009. Hox, Cdx, and anteroposterior patterning in the mouse embryo. *Curr Top Dev Biol* 88:235-255.
- Young T, Rowland JE, van de Ven C, Bialecka M, Novoa A, Carapuco M, van Nes J, de Graaff W, Duluc I, Freund JN, Beck F, Mallo M, Deschamps J. 2009. Cdx and Hox genes differentially regulate posterior axial growth in mammalian embryos. *Dev Cell* 17:516-526.
- Zaret KS, Mango SE. 2016. Pioneer transcription factors, chromatin dynamics, and cell fate control. *Curr Opin Genet Dev* 37:76-81.



Supplemental Figure S1: Comparison of Cdx2 binding in induced EpiSCs and in ESCs overexpressing Cdx2. Upper panel, HoxA cluster; lower panel, HoxB cluster. Cdx2 binding in induced EpiSCs is indicated in red, in ESCs in black. The 3' part of the clusters is bound exclusively in induced EpiSCs, and not in ESCs (region indicated with orange bar). Cdx2 data in ESCs is from Mahony et al., 2014.

A

Gene associated with peak	Position of Cdx2-dependent peak
<i>Hoxa4</i>	(-9318), (-8078), (+869)
<i>Hoxa5</i>	(+3516), (+4756)
<i>Hoxa7</i>	(-4533)
<i>Hoxb4</i>	(-9667), (-138)
<i>Hoxb5</i>	(+4422)
<i>Hoxb7</i>	(-523)
<i>Hoxb8</i>	(-3832)
<i>Hoxb9</i>	(+662)
<i>Hoxc4</i>	(-1525), (+1073), (+8134), (+93920)
<i>Hoxc6</i>	(-16477), (-11305)
<i>Hoxc8</i>	(+2380), (+7552)
<i>Hoxd1</i>	(-41746), (-40965), (-28332), (-15699)
<i>Hoxd3</i>	(-73), (+9305), (+10086), (+22719), (+35352)
<i>Hoxd8</i>	(-3793)



Supplemental Figure S2: Positions activated in a Cdx-dependent way. A) List of *cis*-elements that depend on Cdx for becoming open in the Hox clusters, derived from ATAC-seq data (regions differentially expressed between induced WT cells and induced Cdx null cells). The list includes trunk Hox genes (and distal enhancers computationally linked to *Hoxd1*). B) Upper panel: Stimulation of WT EpiSCs leads to the opening of DNA elements in HoxB at both 3' Hox genes and trunk Hox genes. In Cdx triple null EpiSCs, 3' positions are open as well (orange bars); elements around trunk Hox genes do not become accessible (purple bars). These Cdx-dependent positions overlap with regions that bind Cdx2 in Cdx2-overexpressing ESCs (lower lane, in black). Lower panel: H3K27ac profiles of WT and Cdx null EpiSCs (24 hours of induction) at HoxB; levels of decreased acetylation in Cdx mutant cells correspond to positions that are dependent on Cdx for DNA accessibility.

Chapter 6

General discussion

In this PhD thesis the regulation and function of Hox and Cdx genes are studied in different developmental time frames, like in a trip down the vertebrate antero-posterior body axis. The starting point of the journey is the early moment when the epiblast becomes posteriorized by the expression of *Wnt3*, around embryonic day (E)6.0 of mouse development. At this very start of this journey, in the peri-gastrulation embryo, we explored the *cis*-regulatory landscape and molecular events underlying the earliest Hox gene activation (**Chapter 4**). The local *cis*-conformation of the early Hox landscape and Wnt signaling are important in this phase, together with Wnt-responsive enhancers. As our journey proceeds, the body axis elongates and the Hox genes start to express their middle (or ‘trunk’) members. During this phase Cdx transcription factors become essential players in the posterior genetic network. In the absence of all three Cdx genes, the pool of axial progenitors – including neuromesodermal posterior progenitors (NMPs) – is not maintained and no post-occipital tissues are generated (**Chapter 2**). In addition, the transcription factor T Brachyury is essential as well during this phase: embryos lacking *Cdx2* and *T Brachyury* manifest equally severe truncations as the ‘triple Cdx’ mutants. Moreover we have identified shared direct target genes of these two transcription factors, showing that they collaborate in driving axial elongation (**Chapter 3**). Cdx is not only essential for the maintenance of the NMP niche; it regulates the further colinear activation of the post-initiated Hox clusters (**Chapter 5**). Without Cdx, the *cis*-elements centrally located in the Hox clusters do not become accessible for regulation. Ultimately, at the end of the journey, the 5’-most Hox13 genes arrest axial growth (Young et al., 2009) by competing with Cdx to bind common *cis*-regulatory elements that are important for axial elongation (**Chapter 3**). Thus, Cdx and Hox genes are inherently linked along the entire post-occipital axial elongation.

In this final chapter I concisely discuss the novel findings and major concepts developed in this thesis.

A polarized regulatory landscape and Wnt responsiveness underlie initial Hox activation

Previous research in our lab revealed that the first Hox gene starts to be expressed in the posterior primitive streak region and that this expression is primed one full day before actual transcription (Forlani et al., 2003). However, neither the developmental trigger nor the molecular mechanism underlying initial Hox gene activation had been uncovered.

We have found that Wnt signals act upstream of the initial transcription of the earliest (3') Hox genes, and that this event is immediately followed by a colinear expression of the rest of the cluster members (**Chapter 4**). During this Wnt-induced activation, the Hox clusters become increasingly decorated by active histone modifications and repressive marks are concurrently erased. In our studies we mainly focused on the HoxA cluster, positioned at the border of two large topologically associating domains (TADs) (Dixon et al., 2012). We identified a subdomain of interactions (a 3' subTAD) that comprises the 3' part of the HoxA cluster and the flanking 3' regulatory region which harbors several *cis*-elements ('Ades' enhancers), some of which are Wnt-responsive. These enhancers are all active in the *Hoxa1* endogenous expression domain in the embryo *in vivo*. Noteworthy, the two proximal-most elements (*Ades1* and *Ades2*) are the earliest to be active. Importantly, deletion of these proximal elements results in a decreased response of the 3'-most HoxA gene to Wnt. We conclude that Wnt signals, which are present in the posterior streak at gastrulation stages, are responsible for the spatiotemporal regulation of initial Hox gene expression by acting on a primed 3' *cis*-regulatory landscape.

The trunk Hox genes are dependent on Cdx to be activated

Cdx genes (*Cdx1*, *Cdx2*, and *Cdx4*) have been shown to regulate Hox gene expression domains along the antero-posterior axis (van den Akker et al., 2002; Young and Deschamps, 2009). Our genome-wide Cdx2 binding data revealed that all four Hox clusters are highly occupied by Cdx2 at their 3' and central positions (**Chapter 3**). In order to functionally assess the physiological relevance of this binding, we generated embryo-derived epiblast stem cells (EpiSCs) that lack all Cdx genes. We compared their transcriptional and epigenetic dynamics of Hox activation to the situation in wild type EpiSCs cells (**Chapter 5**). Importantly, Cdx null EpiSCs lose their ability to express Hox genes in a colinear way. Whereas Cdx2 target elements at the 3' side of the cluster could normally become accessible and activate their chromatin, *cis*-elements around trunk Hox genes depend on Cdx to gain DNA accessibility and to become decorated by the active histone mark H3K27ac. These data confirm our previous findings that the 3' part of the Hox clusters are opened and activated by Wnt signals (**Chapter 4**), and demonstrate that Cdx is required for further colinear expression of central Hox genes. At these central positions Cdx acts as a pioneer transcription factor, required to open up the local chromatin.

Wnt-induced and Cdx-dependent Hox regulation is organized in regulatory segments

By taking 4C-seq viewpoints in and around the HoxA locus, we could identify several topological segments that form the architectural blocks within the HoxA cluster (**Chapter 4** and **5**). At the 3' side, a subTAD includes the early regulatory landscape and the 3' part of the HoxA cluster. The middle Hox genes form a second segment, and at the 5' side of the cluster, *Hoxa13* constitutes a distinct entity as it interacts mainly with the 5' TAD (**Chapter 4**). The distribution of *cis*-elements that require Cdx for their opening and activation shows a similar compartmentalization (**Chapter 5**). *Cis*-elements located in the 3' side of the cluster – and in the flanking 3' regulatory region – do not require Cdx for their activation. The domain in which these Cdx-independent elements are located fully overlaps with the topologically defined 3' subTAD. The Cdx-dependent elements are located around the trunk Hox genes and their distribution overlaps with the next topologically defined Hox segment. In **Chapter 4** we have shown that these segments are dynamic during Wnt-induced Hox activation. Altogether, these *cis*-compartments harboring either Wnt-dependent or Cdx-dependent modules are important for, respectively, the initiation and subsequent activation of the anterior and trunk Hox genes. Eventually, the 5'-most topological segment containing the *Hoxa13* gene becomes active and preludes the end of axial growth.

Cdx is essential for the emergence of post-occipital axial tissue

Cdx genes are known to be involved in axial elongation, since embryos carrying mutant alleles of these genes exhibit a shortened body axis (Chawengsaksophak et al., 1997; van den Akker et al., 2002). When all three Cdx genes are inactivated (**Chapter 2**), murine embryos are not able to generate any tissue beyond the occipital level, and axial extension stops after the five anterior-most somites have been generated. The primitive streak, harboring long-term bipotential axial progenitors, loses its activity in the Cdx null mutants and Fgf and Wnt signaling decline. In accordance, culturing *Cdx2* mutant embryos in the presence of exogenous Fgf signals, or overexpressing the Wnt effector *Lef1* in Cdx mutants (Young et al., 2009), both lead to partial rescue of the axial elongation defect. Previously it was demonstrated that similarly rescued phenotypes were obtained by over-expressing the trunk Hox genes *Hoxa5* and *Hoxb8* (Young et al., 2009). Collectively these data indicated that Cdx and central Hox genes constitute a genetic network maintaining trunk progenitor activity by sustaining Wnt and Fgf signaling in the progenitor niche in the posterior growth zone. Different to central and posterior Hox genes that were not activated in Cdx triple mutants, the 3' Hox gene *Hoxb1* was normally transcribed at E7.2.

These observations, added to reported similar findings in other bilaterian animals, point to an ancient role of Cdx in the generation of all tissues posterior to the head. The short-germ band beetle *Tribolium castaneum*, the intermediate-germ band cricket *Gryllus bimaculatus* and the crustacean *Artemia franciscana* all require the Cdx homolog *caudal* for post-head morphogenesis (discussed in **Chapter 1, part A**) (Copf et al., 2004; Shinmyo et

al., 2005). *Cdx/caudal* is thus a master regulator of axial elongation ever since the divergence of the protostome and deuterostome animals approximately 550 million years ago. As could have been surmised from the evolutionary relationship between Hox and ParaHox (*Cdx*) genes, the ancient participation of *Cdx* in differential axial growth is contemporary to the similarly ancient involvement of the Hox genes in axial patterning.

Cdx and T Brachyury transcription factors cooperate in a posterior genetic network

In addition to the *Cdx* and central Hox genes, another important regulator of axial tissue generation is *T Brachyury*, expressed in the posterior growth zone. Mutants lacking this gene exhibit similar truncations as *Cdx2* mutants (Dobrovolskaïa-Zavadskaïa, 1927). Combining the mutant alleles of *Cdx2* and *T Brachyury* leads to a very severe truncation of the body axis, reminiscent to that of the *Cdx* null embryos (**Chapter 3**).

To elucidate whether *Cdx* and *T Brachyury* participate in a common regulatory program, we performed ChIP-seq against these proteins in Wnt and Fgf8-induced EpiSCs that resemble the *in vivo* NMPs (**Chapter 3**). Many common binding sites of *Cdx2* and *T Brachyury* were found in or near Wnt and Fgf signaling pathway components, indicating that these factors co-regulate niche signals in the posterior growth zone. Strikingly, we observed that at some *cis*-elements *Cdx* transcription factors are required for the actual chromatin opening. *Cdx* thus has the ability to function as a pioneer factor, either opening the chromatin itself, or recruiting essential chromatin modifiers to specific target loci.

Epiblast stem cells are an excellent model to study early developmental processes

The transcription of a developmental gene is tightly regulated by *cis*-regulatory elements (i.e. promoters, enhancers, insulators) and the three-dimensional organization of the gene locus (**Chapter 1, part B**). Unfortunately, studying genetic and epigenetic events like transcription factor binding, chromatin opening, histone decoration and chromosomal architecture is very challenging at early embryonic stages. Over the last years several *ex vivo* culture systems have been developed to palliate this difficulty, including the epiblast-derived EpiSCs (Brons et al., 2007; Tesar et al., 2007), gastruloids (van den Brink et al., 2014) and embryonic stem cell (ESC)- or EpiSC-derived neuromesodermal progenitors NMPs [reviewed in (Henrique et al., 2015)]. For stem cell biologists these systems are very useful to examine stem cell capacity in generating differentiated tissues like the motor neurons (Lippmann et al., 2015) or to study the signaling pathways controlling this differentiation (Gouti et al., 2014; Tsakiridis et al., 2015). For developmental biologists eager to study regulation of morphogenesis in a situation relevant for the embryo, these *ex vivo* systems are invaluable to investigate the cellular, genetic and molecular mechanisms that take place during early embryogenesis. Moreover, it is relatively easy to apply gene editing techniques like CRISPR/Cas9 to these models, in order to study the functionality of regulatory elements *in vivo*.

In this thesis we make use of Wnt-stimulated EpiSCs that operate the initial stages of Hox gene activation in a way mimicking the events in early embryo (**Chapter 4**). Upon inducing EpiSCs with Wnt, or Wnt and Fgf8 we also observed activation of the posterior genes *Cdx2* and *T Brachyury* as well (**Chapter 3**). By performing ChIP-seq against these transcription factors in these cells, we could identify their unique and common target loci and unravel their co-operative actions. Cdx null EpiSCs also allowed us to examine the regulatory function of Cdx during trunk Hox gene activation. Because Cdx-deficient embryos are arrested in their axial elongation at the head-trunk transition, they do not express trunk Hox genes as the growth zone is not maintained at post-occipital levels (**Chapter 2**). Therefore, the role of Cdx gene products in regulating trunk Hox genes upon binding, opening and activating the chromatin of central Hox genes would never have been unraveled by studying Cdx mutant embryos exclusively (**Chapter 5**).

Nevertheless, it is essential for developmental biologists that observations made in embryo-derived culture systems like EpiSCs are validated in the embryo. In this thesis we have confirmed the activity of several *cis*-regulatory elements with transgenic *lacZ* reporter embryos (**Chapters 3, 4 and 5**). Moreover, we applied techniques that don't require high number of cells (like RT qPCR, ChIP-qPCR, RNA-seq and ATAC-seq) on early embryonic material to validate regulatory relationships observed in EpiSCs.

In conclusion, we demonstrate in this thesis that the combinatorial use of embryonic material and embryo-derived model systems is very valuable to answer long-standing questions in the field of developmental genetics.

References

- Brons IG, Smithers LE, Trotter MW, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett SK, Clarkson A, Ahrlund-Richter L, Pedersen RA, Vallier L. 2007. Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* 448:191-195.
- Chawengsaksophak K, James R, Hammond VE, Kontgen F, Beck F. 1997. Homeosis and intestinal tumours in Cdx2 mutant mice. *Nature* 386:84-87.
- Copf T, Schroder R, Averof M. 2004. Ancestral role of caudal genes in axis elongation and segmentation. *Proc Natl Acad Sci U S A* 101:17711-17715.
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485:376-380.
- Dobrovolskaia-Zavadskaja N. 1927. Sur la mortification spontanee de la chez la souris nouveau-nee et sur l'existence d'un caractere (facteur) hereditaire, non-viable. *Crit Rev Soc Biol* 97:114-116.
- Forlani S, Lawson KA, Deschamps J. 2003. Acquisition of Hox codes during gastrulation and axial elongation in the mouse embryo. *Development* 130:3807-3819.
- Gouti M, Tsakiridis A, Wymeersch FJ, Huang Y, Kleinjung J, Wilson V, Briscoe J. 2014. In vitro generation of neuromesodermal progenitors reveals distinct roles for wnt signalling in the specification of spinal cord and paraxial mesoderm identity. *PLoS Biol* 12:e1001937.
- Henrique D, Abranches E, Verrier L, Storey KG. 2015. Neuromesodermal progenitors and the making of the spinal cord. *Development* 142:2864-2875.
- Lippmann ES, Williams CE, Ruhl DA, Estevez-Silva MC, Chapman ER, Coon JJ, Ashton RS. 2015. Deterministic HOX patterning in human pluripotent stem cell-derived neuroectoderm. *Stem Cell Reports* 4:632-644.
- Shinmyo Y, Mito T, Matsushita T, Sarashina I, Miyawaki K, Ohuchi H, Noji S. 2005. caudal is required for gnathal and thoracic patterning and for posterior elongation in the intermediate-germband cricket *Gryllus bimaculatus*. *Mech Dev* 122:231-239.
- Tesar PJ, Chenoweth JG, Brook FA, Davies TJ, Evans EP, Mack DL, Gardner RL, McKay RD. 2007. New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* 448:196-199.
- Tsakiridis A, Huang Y, Blin G, Skylaki S, Wymeersch F, Osorno R, Economou C, Karagianni E, Zhao S, Lowell S, Wilson V. 2015. Distinct Wnt-driven primitive streak-like populations reflect in vivo lineage precursors. *Development* 142:809.
- van den Akker E, Forlani S, Chawengsaksophak K, de Graaff W, Beck F, Meyer BI, Deschamps J. 2002. Cdx1 and Cdx2 have overlapping functions in anteroposterior patterning and posterior axis elongation. *Development* 129:2181-2193.
- van den Brink SC, Baillie-Johnson P, Balayo T, Hadjantonakis AK, Nowotschin S, Turner DA, Martinez Arias A. 2014. Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells. *Development* 141:4231-4242.
- Young T, Deschamps J. 2009. Hox, Cdx, and anteroposterior patterning in the mouse embryo. *Curr Top Dev Biol* 88:235-255.
- Young T, Rowland JE, van de Ven C, Bialecka M, Novoa A, Carapuco M, van Nes J, de Graaff W, Duluc I, Freund JN, Beck F, Mallo M, Deschamps J. 2009. Cdx and Hox genes differentially regulate posterior axial growth in mammalian embryos. *Dev Cell* 17:516-526.

Addenda

Nederlandse samenvatting

Acknowledgments / Dankwoord

List of publications

Curriculum vitae

Nederlandse samenvatting

Tijdens de embryonale ontwikkeling van gewervelde dieren (vertebraten), wordt het lichaam geleidelijk van kop (anterieur) tot staart (posterieur) aangelegd. Aan de posterieure zijde van het embryo worden steeds meer structuren gegenereerd, totdat uiteindelijk – in het geval van de meeste gewervelden – de staart is aangemaakt en het volledige lichaam is aangelegd. Bij deze geleidelijke lichaamsas-extentie speelt een populatie stamcel-achtige voorlopercellen een cruciale rol. Deze cellen bevinden zich in het meest posterieure deel van het ontwikkelende embryo, en dragen bij aan weefsels van zowel neurectodermale (zoals ruggenmerg en neuronen) als mesodermale (botten, spieren) herkomst. Door hun bijdrage aan deze twee verschillende weefselstypes, worden deze cellen *neuromesodermal progenitors* (NMP's) genoemd – neuromesodermale voorlopercellen.

In **Hoofdstuk 1 (deel A)** komen de moleculaire en ontwikkelingsbiologische grondbeginselen van de lichaamsas-verlenging aan de orde. Welke transcriptiefactoren zijn betrokken? En welke belangrijke ontwikkelingsbiologische signalen spelen een rol? Ook wordt de lichaamsas-extentie vanuit een evolutionair perspectief behandeld. De geleidelijke groei van kop-tot-staart komt namelijk niet alleen voor bij gewervelden dieren, ook embryo's van geleedpotigen ontwikkelen zich op deze manier. Opvallend daarbij is dat dezelfde transcriptiefactoren en signalen een rol spelen. De manier waarop ons lichaam groeit, is dus evolutionair gezien gebaseerd op zeer oude principes.

Al in een vroeg stadium van de ontwikkeling wordt een anterior-posterieure as in het muizenembryo gedefinieerd. Met andere woorden: aan welke kant van het embryo wordt het hoofd gevormd, en aan welke kant moet de rest van het lichaam worden aangelegd? Hox-genen spelen een belangrijke rol in het identiteit geven aan weefsel dat wordt gegenereerd uit de voorlopercellen: ze dirigeren welke cellen zullen bijdragen aan ribben, of welk type zenuwcel in het ruggenmerg gevormd wordt. Het is daarom niet verrassend dat deze 'identiteits-genen' al snel na de specificatie van de anterior-posterieure lichaamsas worden aangezet.

Hox-genen zijn bijzonder interessante genen om te bestuderen; niet alleen hun functies tijdens de embryogenese, ook hun organisatie in het genoom is zeer opvallend. Er zijn vier Hox clusters (HoxA, HoxB, HoxC, HoxD), die elk bestaan uit ongeveer tien genen. Tijdens de embryonale ontwikkeling wordt elk van de vier clusters geleidelijk geactiveerd, beginnend aan de Hox1-kant (dus voor het HoxA-cluster is dat *Hoxa1*), en gedurende lichaamsas-extentie wordt telkens een opeenvolgend Hox-gen geactiveerd tot het laatste gen op het cluster, Hox13, wordt bereikt. Dit fenomeen van geleidelijke activatie wordt 'Hox colineariteit' genoemd. Het graduele Hox-expressiepatroon dat we tijdens de embryonale ontwikkeling waarnemen, is dus in feite een reflectie van de organisatie van Hox op het chromosoom. In **Hoofdstuk 1 (deel B)** wordt uitgelegd welke factoren – ontwikkelingsbiologische signalen en eiwitten, maar ook de genomische organisatie rondom de Hox-clusters – van belang zijn om dat allereerste Hox1-gen op de juiste plaats en op het juiste tijdstip aan te zetten.

Cdx transcriptie-factoren zijn eiwitten die specifieke DNA-sequenties op het genoom binden, waarbij ze genen reguleren die betrokken zijn bij de generatie van de lichaamsas, maar ook bij het identiteit geven aan de weefsels die aangemaakt worden. In **Hoofdstuk 2** beschrijven we muizenembryo's die alle Cdx-genen (*Cdx1*, *Cdx2*, *Cdx4*) missen. Deze embryo's zijn niet in staat tot lichaams-asverlenging; ze bestaan in feite alleen uit een hoofd. De NMP-cellen die zo belangrijk zijn voor het proces van lichaamsas-verlenging, kunnen in deze embryo's niet in stand worden gehouden. De analyses en experimenten in dit hoofdstuk tonen aan dat Cdx belangrijk is om bepaalde essentiële ontwikkelingsbiologische signalen (Wnt- en Fgf-signalen) te reguleren. Bekende *targets* van Cdx zijn ook de Hox-genen. Een belangrijke observatie is dat de eerste Hox-genen (zoals *Hoxb1*) normaal worden aangeschakeld in embryo's die geen

functionerende Cdx-genen hebben. Latere Hox-genen worden echter nooit geactiveerd – de weefsels waarin dat zou gebeuren worden eveneens niet aangemaakt.

Naast Cdx transcriptiefactoren is er nog een andere zeer belangrijke speler in de embryonale lichaamsas-ontwikkeling: T Brachyury. Zowel Cdx2 (het belangrijkste lid van de Cdx-groep) als T Brachyury is onmisbaar voor de generatie van de lichaamsas. In **Hoofdstuk 3** onderzochten we of deze twee transcriptiefactoren dezelfde genen aansturen tijdens embryogenese. Door gebruik te maken van cellijnen die we van de epiblast van vroege embryos konden maken ('Epiblast-gederiveerde stamcellen', of: EpiSC's), slaagden we erin om met verschillende technieken (ChIP-seq, ATAC-seq) te bestuderen welke DNA-elementen op het genoom door Cdx2 en T Brachyury worden gebonden. We zagen dat er inderdaad veel genen door beide transcriptiefactoren worden aangestuurd: genen van signaalroutes (Wnt, Fgf) die belangrijk zijn voor de instandhouding van voorlopercellen in het embryo.

In **Hoofdstuk 4** laten we zien dat het belangrijke signaleiwit Wnt – dat betrokken is bij de eerste specificatie van de anterieure versus posterieure kant van het embryo – essentieel is voor Hox-activatie. Wanneer EpiSC's in cultuur worden gebracht met Wnt, worden de Hox-clusters gradueel geactiveerd, precies zoals de colineaire expressie in de embryo's. We vonden dat het Wnt-eiwit zijn functie uitoefent op allerlei genetische schakelaars ('enhancers') die dichtbij *Hoxa1* liggen – dus aan de kant van de vroege Hox-genen, waar de colineaire expressie begint. Die enhancers liggen samen met de vroege Hox-genen in een geïsoleerd genomisch segment (een 'subTAD'). Deze subTAD vormt de structurele basis voor het aanzetten van Hox aan de 'goede' kant: specifiek aan deze vroege zijde van het HoxA-cluster wordt het Wnt-signaal ontvangen, waarna allerlei schakelaars eerst worden geopend en vervolgens geactiveerd.

Zoals hierboven genoemd, worden Hox-genen gradueel geactiveerd tijdens embryonale ontwikkeling ('colineariteit'). In **Hoofdstuk 5** duiken we dieper in dit proces en onderzoeken we de rol van Cdx-genen in dit fenomeen. We zagen dat Hox-colineariteit ernstig afweek in EpiSC's die we maakten van embryo's die alle Cdx-genen missen: de genen in de Cdx-mutante cellen worden niet langer keurig één voor één geactiveerd. Het DNA van regulatoire elementen in de Hox-clusters blijkt niet toegankelijk te zijn in Cdx-deficiënte cellen: Cdx is dus belangrijk voor het toegankelijk maken en het activeren van deze DNA-sequenties. Deze afhankelijkheid van Cdx geldt alleen voor de Hox-genen die in het midden van het cluster liggen: de vroege Hox-genen zijn hebben geen Cdx, maar Wnt nodig voor activatie (zie Hoofdstuk 4).

De afhankelijkheid van Wnt voor vroege genen en de afhankelijkheid van Cdx voor midden genen gaat samen met een topologische segmentatie van de Hox-clusters. Zoals boven beschreven liggen de Wnt-afhankelijke vroege HoxA-genen in een subTAD; maar ook de Cdx-afhankelijke genen liggen in een apart segment. Het allerlaatste Hox-gen, *Hoxa13* heeft zijn

eigen segment aan de andere kant van de cluster en is daardoor geïsoleerd van alle andere Hox-genen. We stellen daarom dat de topologische segmentatie van het HoxA-cluster belangrijk is voor het functioneren van Hox. Een dergelijke compartimentalisatie maakt lokale regulatie mogelijk, die alleen effect heeft op genen in hetzelfde segment en niet op genen in een ander segment.

Tot slot worden de belangrijkste bevindingen van dit proefschrift behandeld in **Hoofdstuk 6**.

List of publications

Amin S*, **Neijts R***, Simmini S, van Rooijen C, Tan S, Kester L, van Oudenaarden A, Creyghton MP, Deschamps J. Cdx and T Brachyury co-activate growth signaling in the embryonic axial progenitor niche.

** Authors contributed equally*

In press, Cell Reports

Neijts R, Amin S, van Rooijen C, Tan S, Creyghton MP, de Laat W, Deschamps J. (2016) Polarized regulatory landscape and Wnt responsiveness underlie Hox activation in embryos. *Genes&Development* 30(17):1937–1942

Neijts R, Simmini S, Giuliani F, van Rooijen C, Deschamps J. (2014) Region-specific regulation of posterior axial elongation during vertebrate embryogenesis.

Developmental Dynamics 243(1):88-98

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.

** Authors contributed equally*

Development 139(14):2576-83

van de Ven C, Bialecka M, **Neijts R**, Young T, Rowland JE, Stringer EJ, van Rooijen C, Meijlink F, Nóvoa A, Freund JN, Mallo M, Beck F, Deschamps J. (2011) Concerted involvement of Cdx/Hox genes and Wnt signaling in morphogenesis of the caudal neural tube and cloacal derivatives from the posterior growth zone.

Development 138(16):3451-62

Neijts R*, Amin S*, van Rooijen C, Deschamps J. Cdx is a crucial player during colinear Hox activation and defines a trunk segment in the Hox cluster topology.

** Authors contributed equally*

In revision

Neijts R and Deschamps J. At the start of colinear Hox gene expression: *cis*-features and *trans*-factors orchestrating the initial phases of Hox cluster activation.

Review to appear in a special issue of Developmental Biology (Elsevier) on the centennial of the Hubrecht Institute, (scheduled for 2017)

Other publications:

Jacobs FM, Veenvliet JV, Almirza WH, Hoekstra E, von Oerthel L, van der Linden A, **Neijts R**, Koerkamp MG, van Leenen D, Holstege FC, Burbach JP, Smidt MP. (2011) Retinoic acid-dependent and -independent gene-regulatory pathways of Pitx3 in meso-diencephalic dopaminergic neurons.

Development 138(23):5213-22

Kuribara M, van Bakel NH, Ramekers D, de Gouw D, **Neijts R**, Roubos EW, Scheenen WJ, Martens GJ, Jenks BG. (2012) Gene expression profiling of pituitary melanotrope cells during their physiological activation.

Journal of Cellular Physiology 227(1):288-96

Curriculum vitae

Roel Neijts was born on October 14th 1985 in Milheeze, the Netherlands. In 2004 he received his Gymnasium diploma from the Zwijsen College in Veghel. During his Biology studies at the Radboud University in Nijmegen he performed a short internship at the Department of Cellular Animal Physiology under supervisor of dr. Bruce Jenks. Roel obtained his Bachelor's degree (*cum laude*) in 2009, whereafter he enrolled in the Master's program of 'Cancer Genomics & Developmental Biology' at Utrecht University. For his internship (2009-2010) in the laboratory of dr. Jacqueline Deschamps at the Hubrecht Institute, he received the Wilhelmina Kinderziekenhuis (WKZ) Student Award in 2011. After being intern at the department of Neuroscience and Pharmacology in the laboratory of prof. dr. Marten Smidt (Rudolf Magnus Institute, UMC Utrecht), he received his Master's degree (*cum laude*) in 2011.

In September 2011 he started his PhD research on the function and regulation of Cdx and Hox genes during axial elongation in the laboratory of Jacqueline Deschamps at the Hubrecht Institute. The results of this research are described in this thesis.

Roel obtained a Pasteur-Cantarini fellowship to perform postdoctoral research studies on transcriptional enhancer dynamics in the laboratory of dr. François Spitz in the Pasteur Institute (Paris, France).