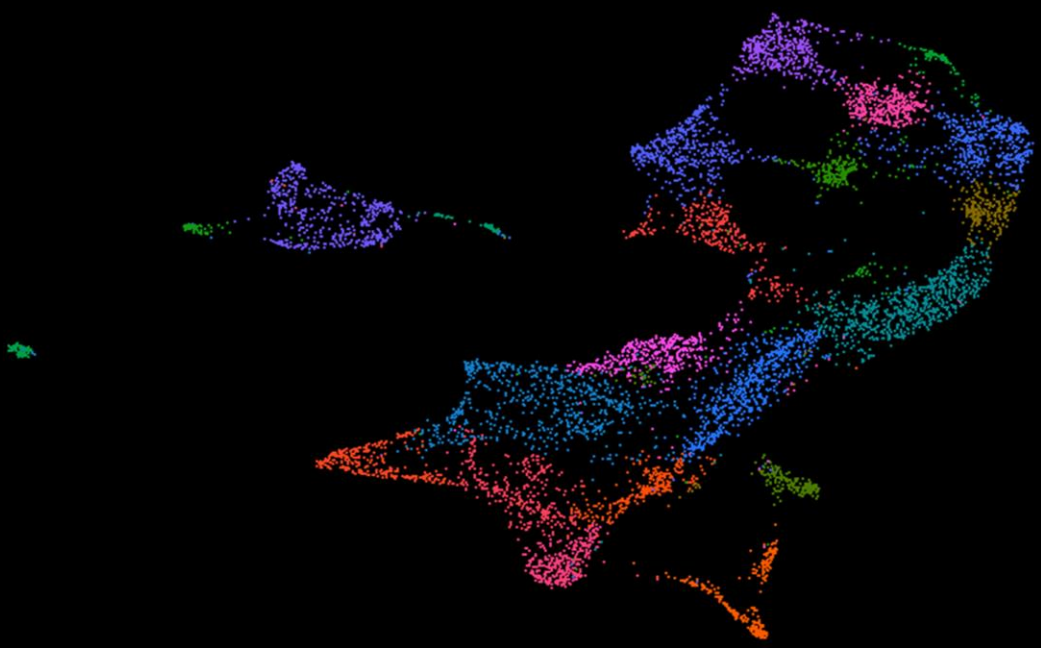
2023

Shifting gears: Insights into the head to trunk transition

Exploring the regulatory changes underlying the head to trunk developmental transition

Patrícia Duarte



Dissertation presented to obtain the **Ph.D degree in**

Integrative Biology and Biomedicine

Oeiras, April, 2023

INSTITUTO DE TECNOLOGIA QUÍMICA E BIOLÓGICA ANTÓNIO XAVIER | UNIVERSIDADE NOVA DE LISBOA

Shifting gears: Insights into the head to trunk transition

Exploring the regulatory changes underlying the head to trunk developmental transition

Patrícia Duarte

Dissertation presented to obtain the Ph.D degree in Integrative Biology and Biomedicine

Instituto de Tecnologia Química e Biológica António Xavier | Universidade Nova De Lisboa

Research work coordinated by: INSTITUTO GULBENKIAN DE CIÊNCIA

Oeiras, April, 2023

This work was supported by the Fundação para a Ciência e a Tecnologia
PD/BD/138240/2018 and COVID/BD152639/2022

Table of Contents

List of Figures

List of Tables

Acknowledgements

To Moisés for the opportunity of joining the lab and working on this project, thank you for your guidance and support. To all the past and present members of the lab I had the privilege of crossing paths with (Ana Casaca, Ana Nóvoa, Anastasiia, André, Artemis, Irma, Kyriel, Luísa, Rita, Tereza and Triin), for all the help and fun times had, all the tea times and banana breaks. To my IBB class, for all the ob/ob dinners and boardgame nights, it was a joy sharing this journey with all of you. To the Collectors and the many adventures in Lanólia. To Cláudia and Alex, for your encouragement, advice, and inspiration in all of these scientific shenanigans. To Pat and Mags, the core team, here through it all.

Abstract

Although vertebrates display a wide diversity of body shapes and sizes their basic body plan is consistently organized into head, neck, trunk and tail structures. During development, the vertebrate main body axis is generated sequentially from head to tail, first by the activity of the primitive streak and later on by the tail bud. Progenitors within these regions, located at the posterior embryonic end, progressively supply cells that will build the different body tissues as the embryo extends along its main body axis. Although this is a continuous process, head and trunk development is controlled by distinct gene regulatory networks that lead to differing cell dynamics and generation of specific tissues.

In the first part of this thesis, we analyze the transcriptomic and chromatin accessibility changes associated with the head to trunk transition in the posterior epiblast region of developing mouse embryos. These analyses revealed changes in WNT signaling, ubiquitination, G-protein signaling and lipid metabolism that may interact to regulate functional properties of the progenitor region of the embryo. In addition, we describe a distinct functional requirement of Wnt palmitoleoylation by Porcn to regulate axial extension and progenitor differentiation during trunk formation. We also show that the head to trunk transition is associated with a vast switch in the regulatory regions modulating either head or trunk development, that impacts the binding activity of key developmental transcription factors.

In the second part, we evaluate the impact of the absence of retinoic acid (RA) signaling in both the transcriptome and chromatin accessibility profiles. We observed that the lack of RA activity had limited influence on chromatin dynamics in the posterior epiblast. In addition, analysis of single cell transcriptomic profiles revealed the disruption of progenitor cell differentiation in the caudal epiblast of *Raldh2* mutants (lacking RA signaling), as well as a reduced spinal cord cell population size. We identified *Fat3* as a potential regulator of this process, as its expression was found downregulated in *Raldh2*

mutants, fitting with its crucial role in neural progenitor proliferation. Throughout this work, we evaluated the functional relevance of various potential enhancers identified in the chromatin accessibility assays. Generation of deletion mutants for each of these elements revealed limited impact on the expression of the target genes, suggesting the existence of redundant enhancers that provide robustness to developmental processes. We also show the existence of regulatory interactions between enhancer elements to produce the proper expression patterns of the target gene. Interestingly, we found that these elements are regulated by different sets of transcription factors.

Overall, we show that the head to trunk transition is associated with vast changes in chromatin accessibility that may be required for the appropriate time- and cell-specific activation of the relevant gene regulatory networks. In addition, the sequencing datasets generated in this thesis provide a valuable resource for future research in embryo development.

Sumário

Apesar dos animais vertebrados apresentarem uma grande variedade de formas e dimensões corporais o plano corporal básico é invariavelmente organizado em cabeça, pescoço, tronco e cauda. Durante o desenvolvimento embrionário, o eixo corporal dos vertebrados é produzido sequencialmente começando na cabeça e terminando na cauda, primeiro pela atividade da linha primitiva e mais tarde pelo botão caudal. À medida que o eixo corporal se vai alongando, os progenitores localizados nestas regiões na parte posterior do embrião, fornecem continuamente células que irão constituir os diversos tecidos axiais. Ainda que seja um processo contínuo, o desenvolvimento da cabeça e do tronco é controlado por redes regulatórias de genes distintas que, por sua vez, conduzem a dinâmicas celulares diferentes e levam à produção de tecidos específicos.

Na primeira parte desta tese, analisamos as mudanças no transcriptoma e na acessibilidade da cromatina associadas com a transição da cabeça para o tronco. Esta análise revelou alterações na sinalização de WNT, ubiquitinação, sinalização mediada por proteína-G e metabolismo de lípidos que poderão interagir de modo a regular propriedades funcionais da região progenitora do embrião. Além disso, descrevemos uma mudança funcional na necessidade de palmitoleoilação de Wnt por parte da proteína Porcn, para regular o alongamento axial e a diferenciação de progenitores durante a formação do tronco. Demonstramos também que a transição da cabeça para o tronco está associada com uma vasta mudança nas regiões regulatórias que controlam o desenvolvimento da cabeça ou do tronco, afetando a dinâmica de ligação de fatores de transcrição fundamentais para o desenvolvimento.

Na segunda parte, avaliamos o impacto da ausência da sinalização do ácido retinóico no transcriptoma e na acessibilidade da cromatina durante o desenvolvimento. A falta de atividade de acido retinóico mostrou um efeito limitado na dinâmica da cromatina na região posterior do epiblasto. A sequenciação do transcriptoma de células individuais revelou perturbações na

diferenciação das células progenitoras do epiblasto caudal de mutantes *Raldh2* (carentes de sinalização do ácido retinóico), assim como redução do número de células da população da medula espinal. O gene *Fat3* foi identificado como um potencial regulador deste fenómeno, uma vez que se encontra reduzida a expressão em mutantes *Raldh2*, concordante com o seu papel crucial na proliferação de progenitores neurais.

Ao longo deste projeto avaliamos a importância funcional de vários potenciais acentuassomos identificados nos ensaios de sequenciação de acessibilidade da cromatina. A obtenção de mutantes com deleções para cada um destes elementos genómicos exibiu um efeito limitado na expressão dos genes sob a sua regulação, o que sugere a presença de acentuassomos redundantes, concedendo robustez aos processos de desenvolvimento embrionário. Observamos também a existência de interações regulatórias entre elementos, essencial para gerar os padrões de expressão corretos do gene sob a sua regulação. Além disso, demonstramos que estes elementos são regulados por diferentes conjuntos de fatores de transcrição.

Em suma, demonstramos que a transição da cabeça para o tronco se encontra associada a extensas mudanças na acessibilidade da cromatina que deverão ser fundamentais para a correta ativação, em termos celulares e temporais, das redes regulatórias de genes relevantes. Por fim, consideramos que os dados de sequenciação gerados nesta tese constituem um importante recurso para estudos futuros dos processos envolvidos no desenvolvimento embrionário.

Chapter 1

Introduction

Early embryonic development

While amniotes share several of the processes regulating their embryonic development, they also show significant differences among clades and even species. I will therefore mostly focus on the features characteristic of early mouse development, as it is the animal model used in the work described in this thesis.

Similarly to other placental mammals, following fertilization, the zygote undergoes a series of cell cleavage divisions, increasing in cell number without altering the overall volume (Aiken et al., 2004). At the implantation stage on embryonic day (E) 4.5, the mouse embryo is composed of: the trophectoderm, an outside layer of cells which will be the precursor of the placenta; the primitive endoderm, that mainly gives rise to the yolk sac; and the pluripotent epiblast, the source of progenitor cells for the future fetus. The epiblast then becomes a cup-shaped epithelium that encloses the proamniotic cavity in a process called cavitation (Bardot and Hadjantonakis, 2020; Coucouvanis and Martin, 1995). At this stage the embryo consists of an epiblast and extraembryonic ectoderm enveloped by a layer of visceral endodermal cells derived from the primitive endoderm. Nodal signaling leads to the specification of the distal visceral endoderm (DVE), a cell population at the distal pole of the embryo (Fig 1.1) (Lu and Robertson, 2004). Bmp4 produced from the extraembryonic ectoderm inhibits DVE formation, limiting this specification to only the distal region of the embryo (Yamamoto et al., 2009). The DVE cells then migrate proximally, breaking the embryo symmetry and specifying the future anterior region of the embryo (Migeotte et al., 2010; Srinivas et al., 2004). The anterior visceral endoderm (AVE) derives from a population of visceral endoderm cells located just caudally of the DVE (Takaoka et al., 2011). The AVE cells produce Lefty1, Cer1 and Dkk1, which respectively inhibit Nodal, BMP and WNT signaling pathways on the prospective anterior region, thus establishing the anterior-posterior axis (Belo et al., 1997; Kemp et al., 2005; Perea-Gomez et al., 2002; Yamamoto et al., 2004). On the opposing side of the epiblast these signaling

pathways remain active, eventually leading to the emergence of the primitive streak at the posterior embryonic end (Perea-Gomez et al., 2002).

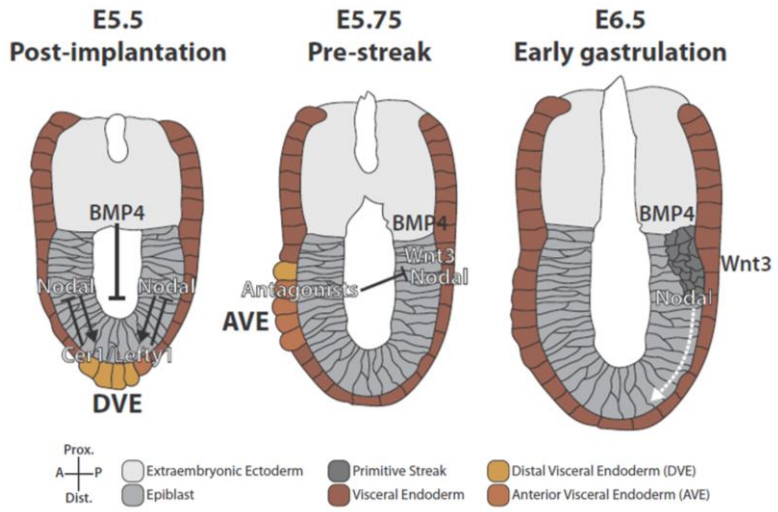


Figure 1.1 – Specification of the visceral endoderm and establishment of the anterior-posterior axis. Nodal signaling results in specification of the DVE. AVE is derived from a visceral endoderm population residing just caudally of the DVE. Secretion of Bmp, Nodal and Wnt antagonists from the AVE restrict the activity of these signaling pathways to the posterior region. The primitive streak emerges at the posterior embryonic end and elongates distally and anteriorly. Adapted from (Bardot and Hadjantonakis, 2020).

Gastrulation starts with the formation of the primitive streak at the posterior end of the embryo. Epiblast cells move towards the midline and undergo an epithelial-to-mesenchymal transition while ingressing through the primitive streak, in a process involving the WNT, BMP, FGF and Nodal signaling pathways (Brennan et al., 2001; Ciruna and Rossant, 2001; Liu et al., 1999; Winnier et al., 1995). As the primitive streak elongates anteriorly, the cells that ingress to the space between the epiblast and the visceral endoderm, form a mesodermal layer that envelops all the embryo (Tam and Behringer, 1997). Meanwhile, the remaining epiblast cells lateral to the primitive streak proliferate to ensure that the pool of cells eventually ingressing through the primitive streak is not prematurely exhausted. Due to the dynamic signaling gradients to which

epiblast cells are exposed, the time and place of ingression along the proximo-distal axis of the primitive streak determines their lineage specification and contributes to the establishment of the body plan (Kinder et al., 1999; Tam and Beddington, 1987). As such, the first cells that ingress at E6.5 from the most posterior part of the epiblast will give rise to the extraembryonic mesoderm. Next, from the posterior region of the epiblast will emerge the cranial and cardiac mesoderm, whereas the cells ingressing through more anterior regions of the primitive streak will generate lateral and paraxial mesoderm. From the node, a specialized structure at the anterior most end of the primitive streak will emerge the definitive endoderm and the axial mesendoderm, which will specify the notochord and prechordal mesoderm (Kinder et al., 1999; Lawson et al., 1991; Quinlan et al., 1995; Tam et al., 1997, 2007).

The epiblast cells anterior to the primitive streak remain within the epithelial layer, where it will form the head ectodermal structures. At around E7.5, this region starts to thicken and subsequently folds, forming the head fold (Fig 1.2) (Downs and Davies, 1993), eventually giving rise to the anterior neural structures that, together with precursors of the cranial mesoderm, definitive endoderm and axial mesendoderm that migrate to this region, generate the primordia of the embryonic head (Sibbritt et al., 2018).

After formation of the head primordia is completed, the embryo starts extending at the posterior end. A large part of this process relies on the neuromesodermal competent (NMC) cell population, capable of generating both neural and mesodermal lineages that progressively supply cells that build the body along the main body axis (Binagui-Casas et al., 2021; Tzouanacou et al., 2009). These progenitor cells are located in the node streak border and in the caudal lateral epiblast and, while some of their derivatives remain in the ectodermal layer and form most of the spinal cord, others gastrulate through the primitive streak to generate paraxial mesodermal structures (Cambray and Wilson, 2007; Wilson et al., 2009). Differentiation of NMC cells into a mesodermal fate requires *Tbxt* and WNT signaling, whereas neural lineage depends on *Sox2* and retinoic acid (RA) signaling (Cunningham et al., 2016;

Gouti et al., 2017; Jurberg et al., 2014; Koch et al., 2017). Moreover, the balance between stem cell renewal and NMC cell differentiation has recently been shown to be regulated by *Sall4* (Tahara et al., 2019). Later in development, with primitive streak regression, which becomes complete around E9.5 (Wilson and Beddington, 1996), NMC cells undergo an incomplete epithelial to mesenchymal transition driven by convergent activity of Snai1 and Tgfbr1 signaling (Dias et al., 2020). NMC cells then relocate to the chordo-neural hinge in the tail bud, where they will provide cells for the formation of tail structures (Tzouanacou et al., 2009).

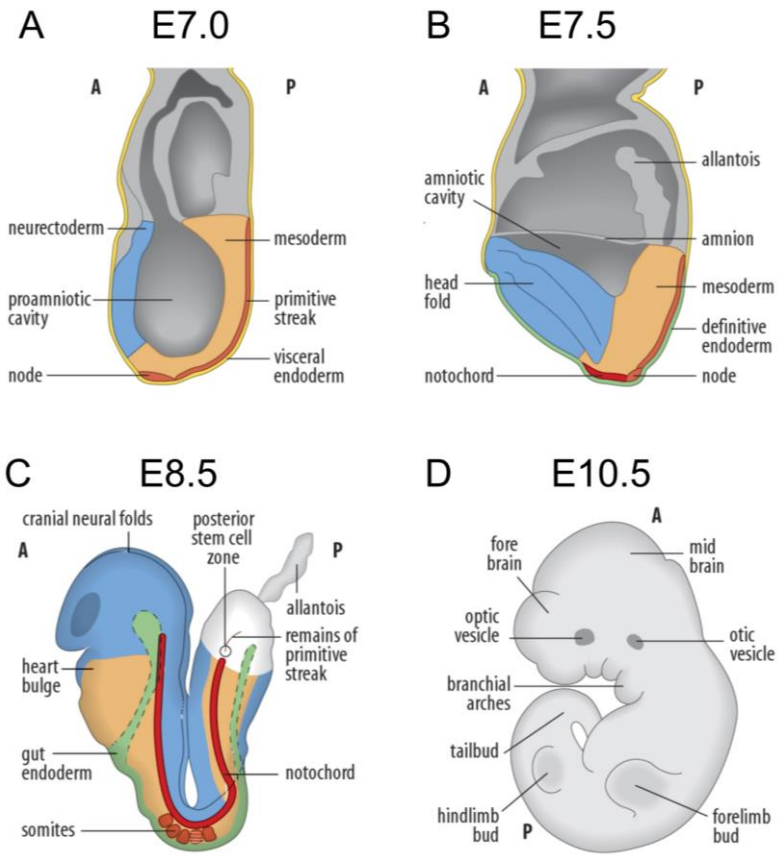


Figure 1.2 - Embryonic development from gastrulation to organogenesis. Schematic representation of mouse embryos at (A) E7.0, (B) E7.5, (C) E8.5 and (D) E10.5 developmental stage. A, Anterior; P, Posterior. Adapted from (Staveley, 2018).

Somitogenesis

The progressive nature of vertebrate axial extension in an anterior-to-posterior direction is most evident in the sequential formation of somites, which give rise to the axial skeleton, skeletal muscle and dermis (Tam, 1981). During somitogenesis, the presomitic mesoderm (PSM) extends posteriorly by addition of new cells from the NMC population and sequentially segments at its anterior end into symmetric units of epithelial tissue, placed on each side of the neural tube (Saga and Takeda, 2001). The periodicity of this process is regulated by the segmentation clock, a network of genes that present a posterior-to-anterior oscillating pattern of expression in the PSM (Palmeirim et al., 1997). As such, with each wave of the segmentation clock a new somite pair is formed at the determination front, a point defined as the region in the PSM where the caudo-rostral gradient of WNT and FGF signaling reaches the threshold level unable to inhibit the segmentation network (Fig 1.3) (Aulehla et al., 2003; Diez del Corral et al., 2003; Dubrulle et al., 2001; Olivera-Martinez and Storey, 2007). While the identity of the oscillatory genes involved in somitogenesis is quite diverse among vertebrates, Notch signaling seems to be a conserved staple of the segmentation clock (Krol et al., 2011). In the mouse, Hes7 and Lfng are involved in the establishment of the segmentation clock oscillations through a negative feedback loop mechanism (Bessho et al., 2001; Morimoto et al., 2005)

Overall, the morphogenic nature of somitogenesis is highly conserved in vertebrates. However, the periodicity of the segmentation is a characteristic feature of each species, ranging from 30 min in zebrafish (Schröter et al., 2008), 90 min in chicken (Palmeirim et al., 1997), 2.5 h in mice (Tam, 1981) to about 5 h in human (William et al., 2007). Moreover, the total number of somites is also species-specific, with 31 somite pairs in zebrafish, 52 in chicken, 65 in mice and about 44 in human (Gomez et al., 2008). Therefore, despite relying on the same process for segmentation, each species can still achieve diversity, which will be reflected in the mature body plan.

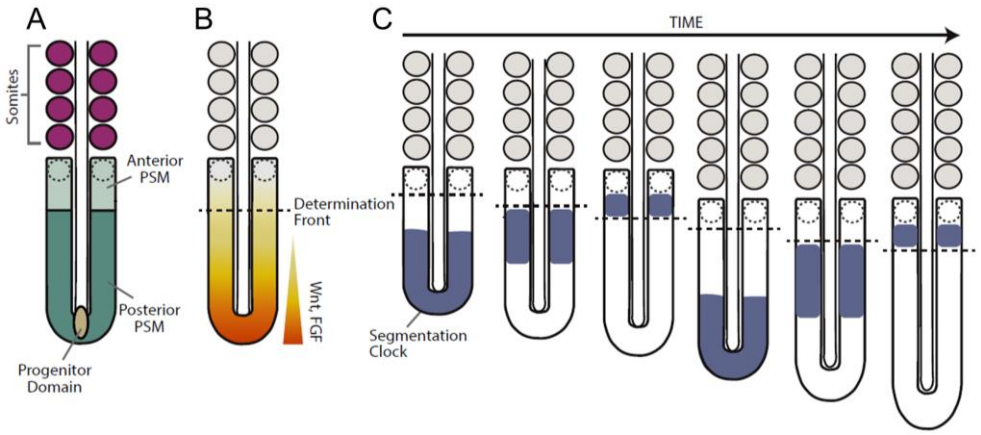


Figure 1.3 – Oscillations of the segmentation clock. (A) Schematic representation of the distinct regions involved in somite formation. (B) The WNT/FGF signaling gradient specifies the position of the determination front in the anterior PSM where the future somite pair (dotted circles) will be formed. (C) Each wave (in blue) of the segmentation clock shows a posterior-to-anterior expression pattern. Adapted from (Diaz-Cuadros and Pourquie, 2021).

Hox genes

The *Hox* genes play a key role in the specification of segment identity along the main body axis, as well as in the morphogenesis of the limbs (Mallo, 2018; Mallo et al., 2010). In mammals they are distributed in four clusters from HoxA to HoxD, originated from two sequential duplication events (Irie et al., 2018), and are divided into 13 paralog groups based on sequence homology and their position within the cluster (Fig 1.4). This results in considerable redundancy among *Hox* paralogous genes (Fromental-Ramain et al., 1996; Horan et al., 1995; Mallo et al., 2010; van den Akker et al., 2001; Wellik and Capecchi, 2003). During development, *Hox* genes are progressively activated from paralog 1 to 13 in a sequence that matches the gene position within the cluster in a 3' to 5' direction, a property called spatial collinearity (Duboule and Dollé, 1989; Graham et al., 1989; Izpisúa-Belmonte et al., 1991). In addition, *Hox* genes also present temporal collinearity, wherein they are progressively activated as the embryo further elongates along the anterior-posterior axis following a sequencing reflecting their organization within the *Hox* cluster (Duboule and Dollé, 1989; Graham et al., 1989; Izpisúa-Belmonte et al., 1991).

For example, *Hox1* genes are involved in proper patterning of the hindbrain (Gavalas et al., 2003); *Hox4* regulate formation of neck vertebrae (Horan et al., 1995); *Hox5*, *6* and *9* are important for ribcage development (McIntyre et al., 2007); *Hox10* determines the lumbar region and *Hox11* specifies the sacral vertebrae (Wellik and Capecchi, 2003). The *Hox13* group genes are the last to become activated and function as growth termination signals of the developing axis (Aires et al., 2019; Economides et al., 2003; Young et al., 2009).

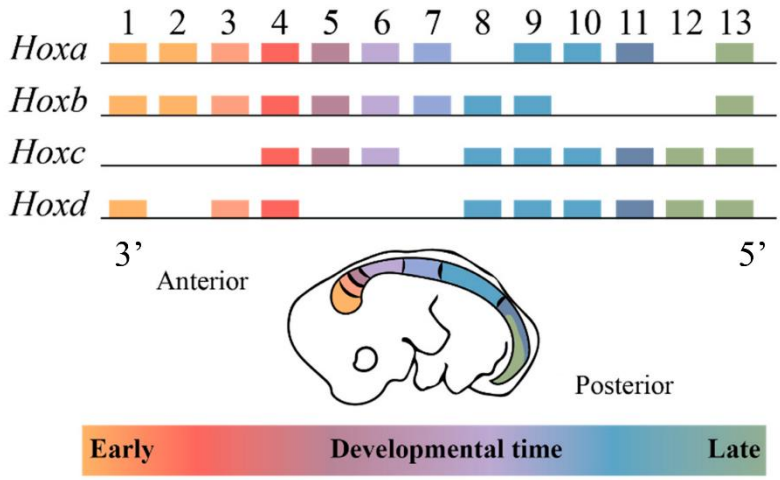


Figure 1.4 – The mammalian *Hox* gene clusters. The Hox clusters present temporal and spatial collinearity, in a way that the timing and pattern of expression of each *Hox* gene matches its position in the cluster. Adapted from (Afzal and Krumlauf, 2022).

Even though *Hox* function has been extensively researched, high sequence homology and overlapping expression patterns still confound proper dissection of individual paralog functions. *Hox* genes code for homeodomain transcription factors that bind to highly similar AT-rich DNA sequences (Berger et al., 2008; Jolma et al., 2013; McGinnis et al., 1984). Interestingly, the participation of other cofactors, in particular members of the TALE family, Pbx and Meis in vertebrates, was revealed to have an important role in increasing Hox transcription factor (TF) binding specificity (Berkes et al., 2004; Jerković et al., 2017; Mariani et al., 2021; Singh et al., 2020).

Despite the role of *Hox* genes in embryonic patterning processes, the establishment of the basic body plan is determined by other regulators acting upstream of *Hox* genes. Gdf11 acting through the Tgfbr1 receptor is a key activator of the trunk to tail transition. Indeed, inactivation of Tgfbr1 blocks this transition (Dias et al., 2020), whereas early activation of this signaling resulted in the formation of the tail at a more anterior position with the concomitant loss of trunk structures (Jurberg et al., 2013). Conversely, *Pou5f1* (also known as *Oct4*) plays a key role in the extension of trunk structures. Conditional inactivation of this gene produced embryos lacking the trunk (DeVeale et al., 2013), whereas sustained expression of *Pou5f1* in mouse embryos led to the extension of the trunk region (Aires et al., 2016). Interestingly, sustained *Pou5f1* expression seems to be the key mechanism producing the remarkably long trunks characteristic of snake species (Aires et al., 2016). In fact, Pou5f1 can regulate *Hox* genes by binding to regulatory regions in the 3' end of the *Hox* cluster that are required for proper gene expression (Tiana et al., 2022). It is still not fully understood how Gdf11/Tgfbr1 activity controls *Hox* genes, but a similar mechanism relying on the control of regulatory regions and chromatin architecture has been suggested (Mallo, 2018).

Head to trunk transition

While vertebrates present a vast diversity of body shapes and sizes, their basic body plan is consistently organized into head, neck, trunk and tail structures. One of the sources of diversity observed in this clade originates from the distinct contribution of each of those regions to the global body plan. For instance, snakes have very long trunks, some species containing more than 300 vertebrae, or in the case of the tail, some lizard species have more than 100 elements, while humans only have a few fused tail vertebrae (Mallo, 2021). Therefore, regulation of the timing of transition between each of these anatomical regions has an important role in establishing this variability. The trunk to tail transition is activated by Gdf11 signaling and associated with a switch from primary to secondary body formation. This is represented by a

change in axial extension, which is no longer reliant on cells ingressing through the primitive streak, depending instead on the activity of the tail bud (Fig 1.5) (Dias et al., 2020; Jurberg et al., 2013; Tzouanacou et al., 2009).

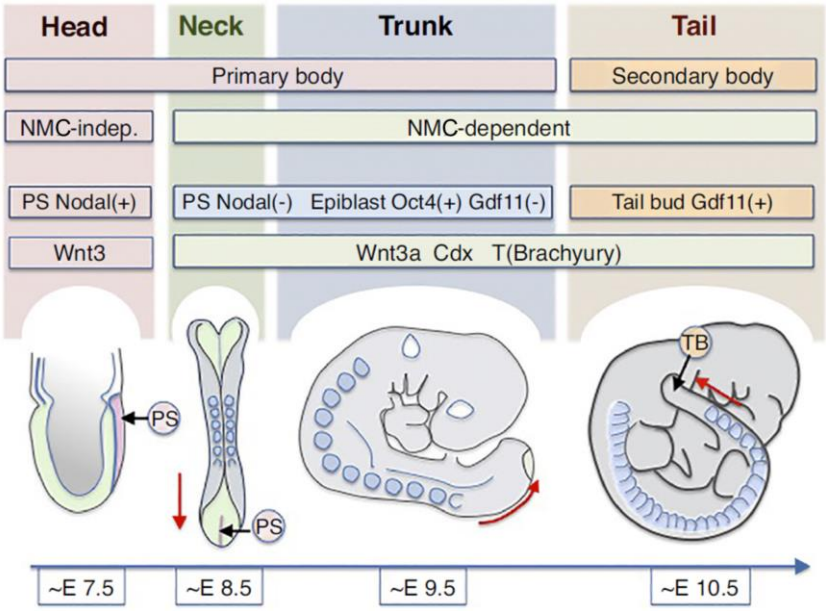


Figure 1.5 – Schematic representation highlighting the reliance on distinct mechanism at different stages of the vertebrate body development. PS, primitive streak; TB, tail bud. The red arrow indicates the direction of axial extension. Adapted from (Aires et al., 2018).

The regulation of the head to trunk transition still poses a challenge but similarly, changes in cellular dynamics and gene regulatory networks involved have also been observed. At the cellular level, the head primordia derive from anterior epiblast tissues and early ingressing cells through the primitive streak prior to the emergence of the NMC population (Sibbritt et al., 2018). The remainder of the body, i.e., the trunk and the tail, relies to a large extent on the activity of the NMC cells during the process of axial extension (Tzouanacou et al., 2009). Regulatory differences have also been identified by genetic approaches. For instance, Nodal signaling only seems to be required for primitive streak activity during early developmental stages (Kumar et al., 2008). Even though later it is crucial for left-right axis specification in the lateral

mesoderm (Brennan et al., 2002), it becomes dispensable for the activity of the primitive streak during post-cranial development (Kumar et al., 2008). Similarly, *Tdgf1* (also known as *Cripto*) regulates cell migration of the DVE and establishment of the AVE, thus playing an essential role in the emergence of the primitive streak, as well as in the formation of mesoderm and endoderm (Ding et al., 1998). In addition, *Eomes* is involved in the specification of the AVE and the epithelial to mesenchymal transition required for primitive streak activity (Arnold et al., 2008; Nowotschin et al., 2013). However, neither of these genes are required later in the regulation of trunk extension.

While head formation requires *Wnt1* and *Wnt3* (Liu et al., 1999; McMahon and Bradley, 1990), trunk development relies on *Wnt5a*, in a partial redundant function with *Wnt11,* and on *Wnt3a*, as well as on key transcription factors like *Tbxt* and the *Cdx* genes (Fig 1.5) (Amin et al., 2016; Andre et al., 2015; Herrmann et al., 1990; Takada et al., 1994; Yamaguchi et al., 1999). This functional specificity is particularly interesting because *Wnt3a*, *Tbxt* and the *Cdx genes* are already expressed at earlier stages but genetic data indicate that they only become functionally relevant during the head to trunk transition. Accordingly, inactivation of these genes originates truncated embryos at the level of this transition (Fig 1.6), whereas the formation of anterior structures remains largely unaffected (Amin et al., 2016; Herrmann et al., 1990; Takada et al., 1994). Therefore, the head to trunk transition might involve a change in the capacity of cells to respond to specific gene networks. Indeed, Tbxt and Cdx2 have been shown to directly bind and activate genes involved in the WNT and FGF signaling pathways during axial extension (Amin et al., 2016).

Absence of RA signaling, obtained by inactivation of *Raldh2*, also leads to developmental arrest at the head to trunk transition (Fig 1.6 D) (Niederreither et al., 1999). Notably, acute exogenous administration of RA at the stage of transition can rescue trunk development (Zhao et al., 2009), suggesting that unlike the previous examples, RA signaling may only be required during the transition phase.

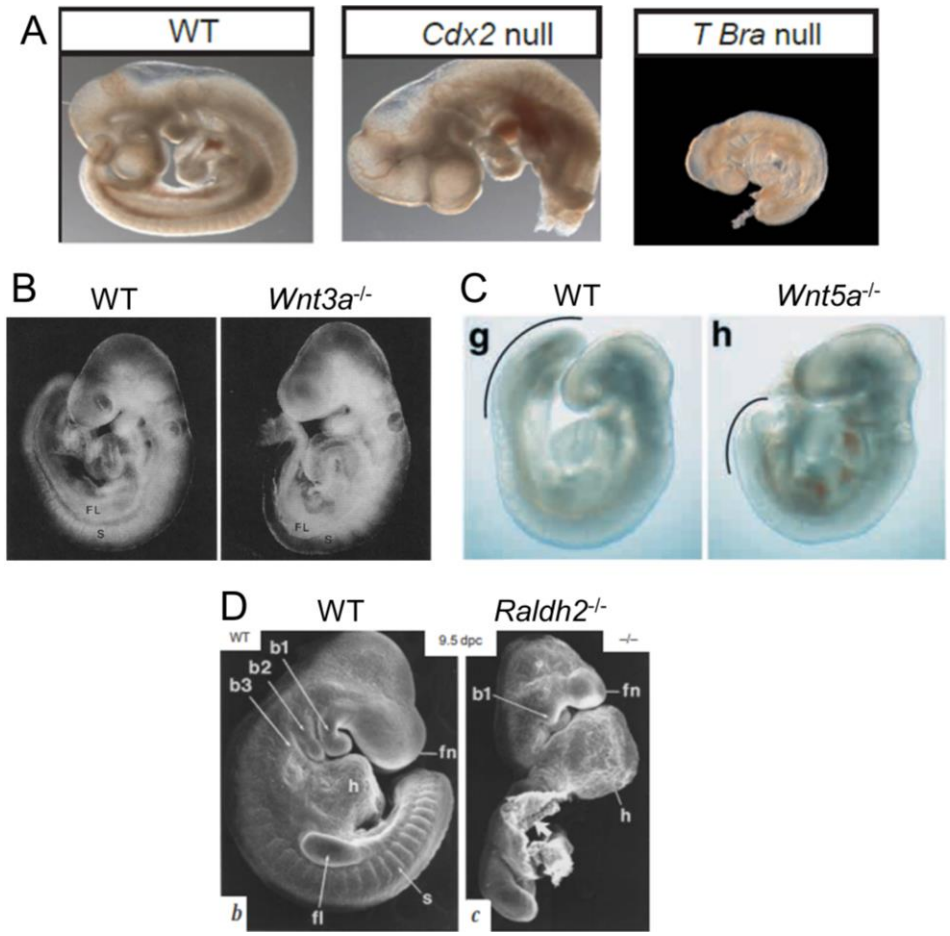


Figure 1.6 – Mutant embryos displaying posterior truncations. (A) *Cdx2*$^{-/-}$ and *Tbxt*$^{-/-}$ mutant embryos at E9.5. Adapted from (Amin et al., 2016). (B) *Wnt3a*$^{-/-}$ mutant embryos at E9.5. Adapted from (Takada et al., 1994). (C) *Wnt5a*$^{-/-}$ mutant embryos at E9.5. Adapted from (Yamaguchi et al., 1999). (D) *Raldh2*$^{-/-}$ mutant embryos at E9.5. Adapted from (Niederreither et al., 1999).

WNT signaling

The WNT signaling pathway is involved in many cellular functions, controlling a wide range of processes such as cell proliferation, differentiation, cell adhesion, cytoskeletal rearrangement, migration and patterning (Logan and Nusse, 2004; Takada et al., 1994; Wang et al., 2010; Yamaguchi et al., 1999). WNT signaling is of particular interest in the context of embryonic development, as shown by the lethal phenotypes observed in knockout mutants for several members of this family. For example, *Wnt1* homozygous mutants

lack midbrain and cerebellum structures (McMahon and Bradley, 1990); mutants for *Wnt3a* exhibit a truncated axis due to an imbalance of progenitor differentiation to a neural fate, at the expense of mesodermal tissues (Takada et al., 1994); and *Wnt5a* mutants also display axial defects, in this case due to the role of *Wnt5a* in regulating the proliferation of progenitor cells (Yamaguchi et al., 1999).

Wnt signals can operate through different pathways. The canonical pathway involves regulation of β-catenin. In the absence of Wnt activation, β-catenin is continuously phosphorylated by Gsk3, which together with Apc, Axin and Ck1α form the destruction complex (Fig 1.7) (Ikeda et al., 1998; Liu et al., 2002; Siegfried et al., 1992). This phosphorylation allows Btrc, an E3 ubiquitin ligase, to recognize and ubiquitylate β-catenin, leading to its subsequent proteasomal degradation (Aberle et al., 1997; Kitagawa and Hatakeyama, 1999; Winston et al., 1999). WNT signaling is activated by binding of Wnt ligand to a member of the Fzd membrane receptor family, which forms a complex with coreceptor Lrp5 or Lrp6, triggering signal transduction (Bhanot et al., 1996; Cong et al., 2004; Pinson et al., 2000). This prompts the recruitment of Dvl to the membrane and subsequent interaction with the destruction complex (Bilic et al., 2007). This results in the release of β-catenin from the destruction complex, its accumulation in the cytoplasm and subsequent transport to the nucleus, where it interacts with the Tcf/Lef family of transcription factors to activate target gene expression (Behrens et al., 1996; Hernández et al., 2012; Wetering et al., 1997).

In addition to the β-catenin dependent canonical pathway, Wnt ligands can act through non-canonical pathways, including the WNT/Planar cell polarity or the WNT/Ca^{2+} pathways. These cascades are not mediated by β-catenin and, instead of relying on Lrp5/6, Fzd receptors interact with coreceptors Ror/Ryk (Oishi et al., 2003; Yoshikawa et al., 2003). The WNT/Planar cell polarity pathway signals via small GTPases and Jnk to coordinate cytoskeletal reorganization and cell motility (Boutros et al., 1998; Fanto et al., 2000; Oishi et al., 2003; Strutt et al., 1997). The Wnt/Ca^{2+} pathway signals through

heterotrimeric G-proteins to activate phospholipase C, which in turn leads to the release of Ca^{2+} from the endoplasmic reticulum and consequent activation of Ca^{2+} dependent signaling (Kühl et al., 2000; Slusarski et al., 1997; D. C. Slusarski et al., 1997).

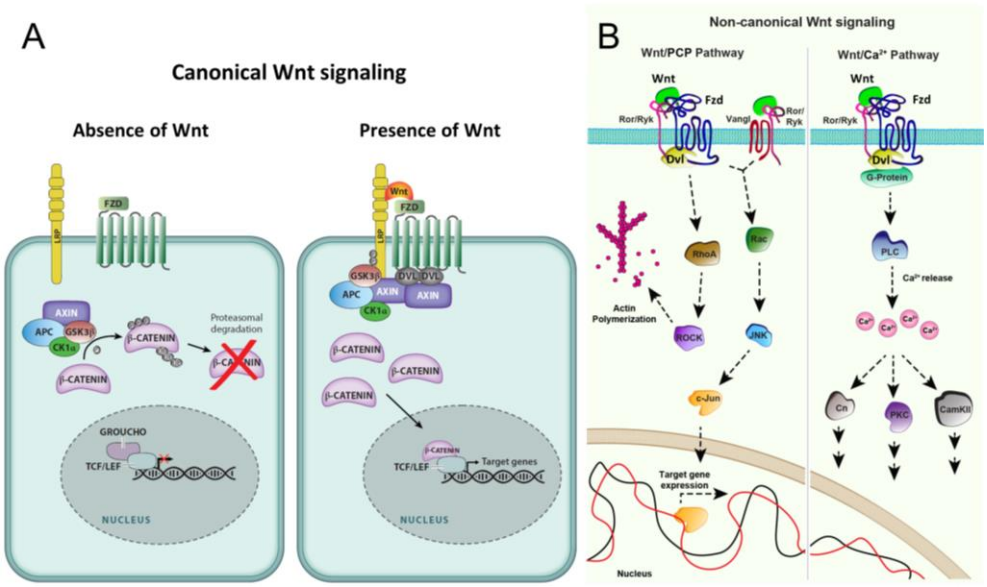


Figure 1.7 – Wnt signaling pathways. (A) In the absence of Wnt ligand, β-catenin is phosphorylated by the destruction complex, leading to its ubiquitination and proteasomal degradation. In the nucleus, the transcriptional repressor Groucho prevents Wnt target gene expression. Upon binding of Wnt ligand, the destruction complex is recruited to the membrane, resulting in the accumulation of β-catenin, which is then translocated into the nucleus where it interacts with Tcf/Lef to activate expression of target genes. Adapted from (Rim et al., 2022). (B) Wnt ligands can also trigger the non-canonical signaling pathways through the coreceptors Ror/Ryk. In the Wnt/Planar cell polarity pathway, activation of small GTPases RhoA and Rac lead to actin polymerization and activation of the Jnk pathway. In the Wnt/Ca^{2+} pathway, G-protein signaling activates Plc thus increasing the intracellular levels of Ca^{2+} and triggering Ca^{2+} signaling. Adapted from (Mehta et al., 2021).

Regardless of the pathway they will activate, palmitoleoylation of Wnt proteins by Porcn is thought to be required for their binding to Wls, which mediates the transport of Wnt proteins from the endoplasmic reticulum to the plasma membrane, where they are secreted to the extracellular space

(Bänziger et al., 2006; Coombs et al., 2010; Herr and Basler, 2012; Rios-Esteves et al., 2014; Takada et al., 2006). In the absence of Porcn activity Wnt proteins (most prominently Wnt3) accumulate in the endoplasmic reticulum and embryonic development arrests at the gastrulation stage (Biechele et al., 2011; Takada et al., 2006).

Retinoic acid signaling

Signaling by retinoic acid (RA) involves two types of receptors, the RAR and the RXR, each with different isoforms (Benbrook et al., 1988; Brand et al., 1988; Dawson and Xia, 2012; Giguere et al., 1987; Zelent et al., 1989). In the absence of RA, RAR and RXR form heterodimers that bind to retinoic acid response elements (RAREs) and interact with nuclear corepressors, Ncor1 and Ncor2 (Chen and Evans, 1995; Kurokawa et al., 1995). These corepressors recruit histone deacetylases (HDAC) and the Polycomb repressive complex 2 (PRC2), leading to deposition of H3K27me3 and consequent folding of the chromatin into a closed state (Hong et al., 2018). The binding of RA to RAR/RXR dimers induces a conformational change resulting in the release of the corepressors and consequent recruitment of coactivators, like Ncoa and histone acetylases that increase chromatin accessibility and gene activation (Dey et al., 2007). RA signaling can also lead to repressive functions, as is the case for the RARE located upstream of *Fgf8*, in which Ncor functions ligand-dependently to recruit HDAC and PRC2 (Kumar et al., 2016; Kumar and Duester, 2014). Nevertheless, it is still not understood what distinguishes activating from repressive RAREs.

Retinoic acid is metabolized from vitamin A (retinol), first by the action of retinol dehydrogenase (Rdh10) to produce retinaldehyde (Sandell et al., 2012), followed by conversion to retinoic acid by tissue specific retinaldehyde dehydrogenases, Raldh1, Raldh2 and Raldh3 (Dupé et al., 2003; Fan et al., 2003; Niederreither et al., 1999). While *Raldh1* and *Raldh3* expression is most prominent in the optic vesicles (Dupé et al., 2003; Fan et al., 2003), at E8.5 *Raldh2* is expressed in the somites, lateral mesoderm, heart mesoderm and

the optic vesicles (Niederreither et al., 1997). The synthetized RA can then participate in intracellular signaling or be secreted, triggering paracrine signaling. Cyp26a1 and other enzymes of the P450 family, have an important role in the regulation of tissue specific RA signaling by degrading RA molecules (Hernandez et al., 2007; Sakai et al., 2001). This regulation is particularly crucial in the establishment of RA signaling gradients in the developing trunk, as Cyp26a1 expressed in the tailbud degrades RA coming from more anterior tissues (Fig 1.8) (Sakai et al., 2001).

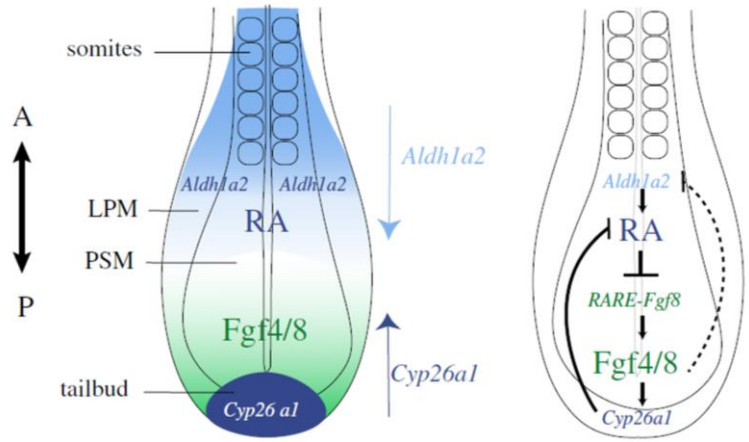


Figure 1.8 – RA signaling in the mouse posterior epiblast at E8.5. Diagrams detailing the opposing gradients of RA (blue) and Fgf (green) along the anterior-posterior axis. Cyp26a1 is expressed in the tailbud and inhibits RA signaling by degrading RA molecules. A, anterior; LPM, lateral plate mesoderm; P, posterior; PSM, presomitic mesoderm. Adapted from (Bernheim and Meilhac, 2020).

RA signaling has been implicated in the development of the eye, hindbrain, heart, trunk and limbs, as well as in spermatogenesis (Busada et al., 2015; Hernandez et al., 2007; Matt et al., 2005; Niederreither et al., 2001, 1999; Zhao et al., 2009). Although some of the genes under RA regulation are known, further studies are required to unveil the mechanisms of RA signaling in specific tissues and developmental stages. Mice lacking *Raldh2* display problems in trunk somitogenesis and premature arrest of body axis extension, failing to undergo axial rotation and heart tube looping (Fig 1.6 C) (Niederreither et al.,

2001, 1999). These mutants also present smaller somites due to lack of RA-mediated repression of *Fgf8* expression (Fig 1.8), thereby expanding the gradient of Fgf8 along the PSM to control somite size (Cunningham et al., 2015a). Although still not fully understood, axial truncation may be linked to the altered balance in NMC differentiation observed in *Raldh2* mutants, owing to decreased *Sox2* expression in prospective neural progenitors and the increase of mesodermal progenitors expressing *Tbx6* (Cunningham et al., 2015a).

Chromatin structure and organization

In eukaryotic cells, the genome is packed in the nucleus in the form of chromatin, a complex consisting of DNA, RNA, histones, and non-histone proteins. The core structural unit of the chromatin is the nucleosome, which is composed of 147 base pairs of DNA wrapped around an octamer comprised of two of each of the core histone proteins H2A, H2B, H3, and H4 (Bednar et al., 1998; Luger et al., 1997). Nucleosomes are connected by linker DNA, associated with a H1 histone (Allan et al., 1980; Bednar et al., 1998).

The existence of organized chromosome territories has long been described, initially distinguishing euchromatin (active and open regions) from heterochromatin (repressed and closed regions) (Heitz, 1928). More recently, development of high-throughput chromosome conformation capture (Hi-C) methods (Lieberman-Aiden et al., 2009) has provided ever increasing details into the fine structure of the genome organization within the nucleus. The resource to such methods revealed the existence of distinct A and B chromatin compartments, representing respectively active and inactive chromatin (Lieberman-Aiden et al., 2009). In addition, at higher resolution, the chromatin is organized in preferentially interacting domains known as topologically associating domains (TADs), DNA regions with a size range of ~100 kb to ~1 Mb (Dixon et al., 2012). Chromatin interactions are established at high frequency within TADs, whereas very few contacts are established with regions outside their boundaries (Dixon et al., 2012). In vertebrate cells TAD boundaries are enriched in CTCF and cohesin. TADs are thought to be

produced through a loop-extrusion process, in which DNA slides through a ring-shaped cohesin molecule attached to a CTCF site until the appearance of another CTCF site in a convergent orientation, leading to stabilization of the cohesin-CTCF interaction and creating an insulated looped domain (Rao et al., 2014; Zuin et al., 2014). Transcriptional regulation often occurs within TADs, as they facilitate bringing distal regulatory elements into close proximity to their target genes (Fig 1.9). Disruption of TAD boundaries often result in the establishment of ectopic chromosomal contacts that can lead to transcriptional misregulation (Lupiáñez et al., 2016; Narendra et al., 2015; Nora et al., 2012). Although TADs were initially considered to be stable structures that are mostly conserved between tissues and across vertebrate species (Dixon et al., 2012; Nora et al., 2012; Woltering et al., 2014), recent studies have shown that TADs can be dynamic structures, acquiring cell state or tissue specific configurations (Gabriele et al., 2022; McArthur and Capra, 2021).

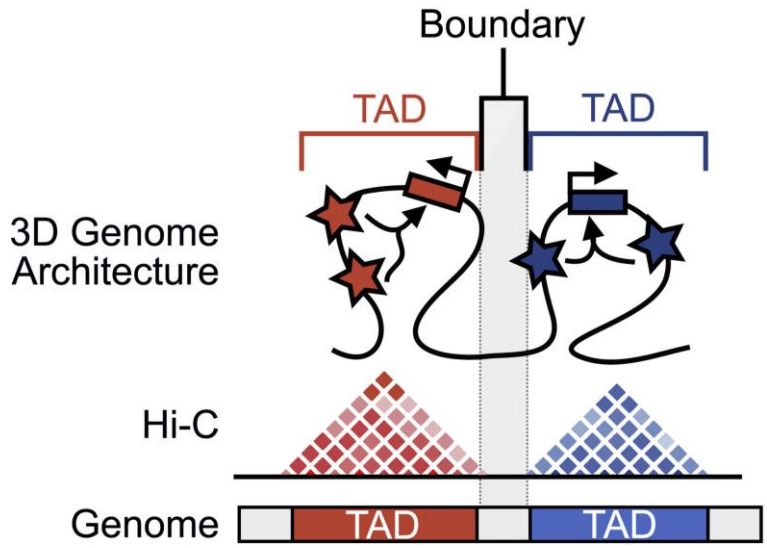


Figure 1.9 – Schematic representation of TAD organization. Regions within a specific TAD preferentially interact with one another than with regions outside, as represented by the Hi-C interaction heatmaps. Stars represent regulatory regions and boxes depict target genes. Adapted from (McArthur and Capra 2021).

Chromatin accessibility

Most processes involving genomic DNA, including transcription, replication or DNA repair require accessibility of the relevant factors to the DNA molecule. Nucleosome occupancy and positioning thus play an important role in the regulation of those processes by modulating the access and occupancy of the cellular machinery to DNA through steric occlusion (Widom, 1998). Nucleosome occupancy is dynamically regulated throughout the genome, with active promoters and enhancers showing higher nucleosome turnover rates than inactive or heterochromatin regions (Deal et al. 2010; Schones et al. 2008). This dynamic regulation can be achieved through the action of chromatin remodelers, TFs, or by RNA polymerase in actively transcribing genes (Clark and Felsenfeld, 1992; Iwafuchi-Doi et al., 2016; Wang et al., 2009).

Post-translational modifications of histone N-tails are another important factor determining chromatin structure and include phosphorylation, acetylation, ubiquitylation, and methylation, among others. These modifications function following two general mechanisms, namely by recruiting architectural non-histone proteins that can alter chromatin conformation (Bannister et al., 2001), or by directly altering protein charge (Hebbes et al., 1994). Histone acetylation neutralizes the charge of lysine residues, thus promoting chromatin unfolding (Hebbes et al., 1994). In general, high levels of histone acetylation are associated with active chromatin. Histone tail acetylation marks have been mapped to promoters and transcribed regions of active genes (H3K4ac), as well as regions surrounding Transcription Start Sites (TSS) (H3K27ac) (Wang et al., 2008). On the other hand, histone methylation appears to be more versatile, being involved in both gene activation and repression. Histone methylation is characterized by introduction of methyl groups, mainly to lysine residues of histone H3, and can occur in either a mono-, di-, or tri-methylated form (Murray 1964; Paik and Kim 1967; Hempel et al. 1968). This diversity allows for a wide range of effects on gene activity. For instance, H3K4me3 is highly enriched at active promoters near TSS (Bernstein et al., 2002).

Conversely, H3K9me3 and H3K27me3 are associated with repressive functions. H3K9me3 is characteristic of heterochromatin while H3K27me3 marks dynamically regulated genes (Barski et al., 2007; Peters et al., 2003).

During development, genes can be decorated with bivalent histone marks, meaning the simultaneous presence of both activating H3K4me3 and repressive H3K27me3 marks. This pattern is typically found in earlier developmental stages before lineage commitment (Bernstein et al., 2006). As differentiation progresses and lineages are specified, one of the two marks is lost, thus leading to appropriate gene activation or repression. An example of this process is provided by the *HoxD* cluster, where this initial bivalent chromatin state is resolved into a distinct distribution of histone marks along the cluster. Actively transcribed genes are labelled with H3K4me3, whereas inactive genes are only covered with H3K27me3 (Soshnikova and Duboule, 2009). Interestingly, the presence of bivalent histone marks has also been described in enhancers. Active enhancers bear H3K4me1 together with H3K27ac while poised enhancers display H3K4me1 with H3K27me3, and primed enhancers retain H3K4me1 but lack any H3K27 modification (Creyghton et al., 2010; Rada-Iglesias et al., 2011). Even though both poised and primed enhancer states are unable to drive gene expression, the transition between these states allows dynamic regulation of enhancer activity.

Regions with an open chromatin conformation display nucleosome displacement, specific histone modification signatures and are DNase I hypersensitive sites. This allowed the first forays into genome-scale measurements of open chromatin, by mapping DNase I hypersensitive sites using microarrays (Crawford et al., 2006; Sabo et al., 2006). Henceforth, several other techniques were developed to evaluate chromatin accessibility, all relying on the basic principle that open regions of the chromatin are more susceptible to enzymatic cleavage (DNase-seq, MNase-seq, ATAC-seq) (Boyle et al., 2008; Buenrostro et al., 2013; Schones et al., 2008), methylation (NOME-seq) (Han et al., 2011) or mechanical shearing (FAIRE-seq) (Giresi et al., 2007). Among them, ATAC-seq (Assay for Transposase Accessible

Chromatin using sequencing) has rapidly emerged as one of the most widely adopted approaches to profile chromatin accessibility because of its simplicity, high sensitivity and the possibility of downscaling to single cell resolution (Buenrostro et al., 2015; Corces et al., 2017; Cusanovich et al., 2015). ATAC-seq uses a hyperactive Tn5 transposase, which simultaneously cuts DNA and inserts sequencing adaptors at regions of increased accessibility that are then sequenced and mapped to the reference genome (Fig 1.10) (Buenrostro et al., 2013).

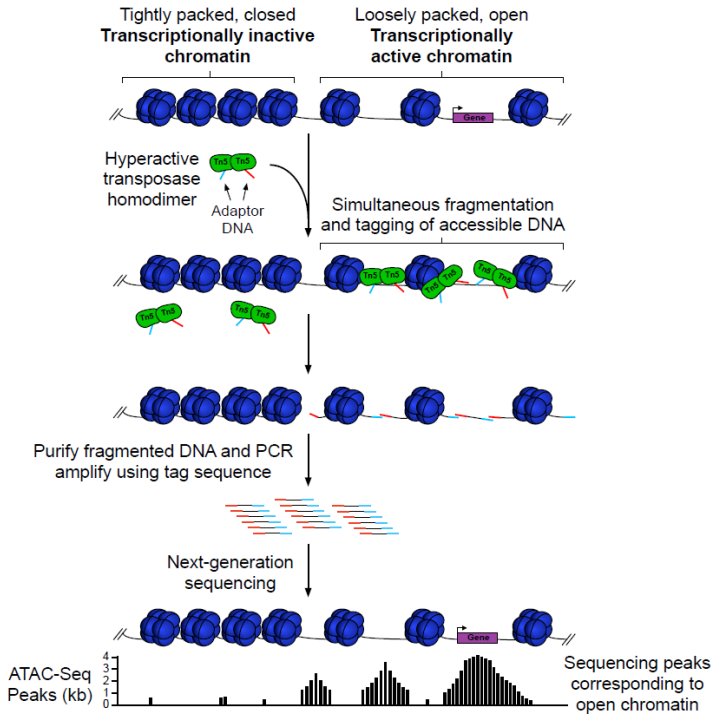


Figure 1.10 – Schematic representation of the molecular basis of ATAC-seq. In ATAC-seq, hyperactive Tn5 Transposase is pre-bound to synthetic adaptor sequences, the resulting transposase activity leads to fragmentation of DNA and end-joining of the tagged adaptor sequences. DNA libraries are obtained by PCR amplification with tag specific primers. Finally, the resulting sequencing reads are mapped to an assembled genome to identify the regions with higher accessibility. Adapted from (Deema, 2019).

Enhancer Regulation

A hallmark of development is the progressive emergence of distinct gene regulatory networks that specify identity as cells differentiate. Regulation of such transcriptional programs relies on the dynamic interaction between regulatory elements and their target genes in a time- and cell-specific manner. Enhancers were originally identified as regulatory DNA regions that activate gene expression regardless of orientation and position relative to their target gene (Banerji et al., 1981). Although the mechanisms of how enhancers function to activate promoters are still not completely understood, in the current view TFs bind to enhancers recruiting other TFs, coactivators, chromatin remodelers, the Mediator complex and RNA Polymerase II to the promoter of the target gene, leading to transcription initiation and elongation (Lee et al., 2015; Phillips-Cremins et al., 2013; Sawado et al., 2003). While promoters are found in close proximity and upstream of the TSS of the respective gene, enhancers can be located both at close or at considerable distance from their target genes. Moreover, enhancers do not necessarily interact with the closest gene. The ZRS, an enhancer that activates *Shh* in the developing limb bud, is a paradigmatic example of this characteristic, as it is located within intron 5 of the *Lmbr1* gene at about 1Mb downstream from the *Shh* transcription unit (Lettice et al., 2003). This feature further complicates efforts to associate enhancers to their target genes. Since enhancers can be situated quite far away from their targets, looping of the chromatin is required to bring enhancers into close proximity with the respective promoters. As such, chromosome conformation capturing techniques like 4-C, Hi-C and ChIA-PET, have been instrumental to identify enhancer-promoter interactions (Kieffer-Kwon et al., 2013; Mifsud et al., 2015).

Thesis aims

The main objective of this project was to explore the mechanisms behind the switch in cell competence during head to trunk transition and the role that retinoic acid plays in this process. Most particularly, we aimed to:

- Explore gene regulatory networks involved in the head to trunk transition.
- Uncover genomic areas involved in this developmental transition and their role in regulating gene expression.
- Assess the impact of RA signaling on chromatin accessibility and gene expression.

Chapter 2

Regulatory changes associated with the head to trunk developmental transition

Part of chapters 2 and 3 are included in:

Duarte P., Correia R., Nóvoa A., Mallo M. Regulatory changes associated with the head to trunk developmental transition (2022) bioRxiv. doi.org/10.1101/2022.12.18.520961

Author contributions

Patrícia Duarte and Moisés Mallo designed experiments. Patrícia Duarte performed experiments. Ana Nóvoa performed pronuclear microinjection of DNA constructs. Rion Correia generated the protein-protein interaction network. Patrícia Duarte and Moisés Mallo analyzed data. Patrícia Duarte wrote this chapter and Moisés Mallo edited it.

Introduction

During embryonic development the vertebrate body is generated progressively in a head to tail sequence. Although this is a continuous process it occurs in distinct steps that produce head, neck, trunk and tail structures (Aires et al., 2018; Stern et al., 2006; Wilson et al., 2009). Each of these stages is characterized by distinct cell dynamics and the generation of a specific set of tissues. For instance, during head development, the embryo establishes the main body axis, lays down the anlage for future brain structures and engages in the process of gastrulation to generate the germ layers (Stern et al., 2006; Tam and Behringer, 1997). The latter process requires the induction of the primitive streak at the posterior end of the embryo that organizes the emergence of the embryonic endoderm as well as the mesodermal tissues for the head and heart primordia (Tam and Behringer, 1997). Genetic analyses in mice have identified key regulators involved in these processes. Some examples include interactions between *Nodal*, *Bmp4* and *Wnt3* to form the primitive streak (Bardot and Hadjantonakis, 2020), *Eomes* for the specification of the endodermal layer and mesoderm delamination (Arnold et al., 2008), and *Gata4* and *Gata6* for heart induction (Zhao et al., 2008). The switch to trunk development is associated with major changes in the growth dynamics of the embryo. It starts elongating the main body axis at the posterior embryonic end by the progressive addition of new tissue produced by the activity of axial progenitors (Aires et al., 2018; Steventon and Martinez-Arias, 2017; Wilson et al., 2009). This process is associated with the emergence of the neuro-mesodermal competent (NMC) cell population, the progenitor cells that build the spinal cord and the axial skeleton (Aires et al., 2018; Binagui-Casas et al., 2021; Steventon and Martinez-Arias, 2017; Wymeersch et al., 2021). Additional progenitors in the epiblast also lay down the tissues that will contribute to the formation and vascularization of the organs involved in digestive, excretory and reproductive functions of the animal (Ferretti and Hadjantonakis, 2019). Similarly to the cells contributing to most embryonic tissues during head

development, the progenitors generating trunk structures are also part of the epiblast, which at this stage occupies the posterior end of the embryo (Aires et al., 2018; Wymeersch et al., 2021, 2016). Also, the primitive streak keeps being the main organizer of progenitor activity during trunk development (Aires et al., 2018; Steventon and Martinez-Arias, 2017). However, the regulatory processes undergo major changes. Inactivation of *Tbxt*, the *Cdx* genes, *Wnt3a*, and the combined *Wnt5a* and *Wnt11* loss of function results in embryo truncation at the head to trunk transition, indicating their essential role for trunk development (Andre et al., 2015; Chawengsaksophak et al., 1997; Chesley, 1935; Herrmann et al., 1990; Savory et al., 2011; Takada et al., 1994; Yamaguchi et al., 1999).

These observations indicate that the transition into trunk development is associated with a global change in gene regulatory networks, most particularly in the posterior region of the embryo, that switches from gastrulation movements to axial extension. Importantly, many of the factors that control developmental processes during trunk extension are also expressed at earlier stages of development, despite not being required at those stages according to genetic experiments. This indicates that the head to trunk transition also involves a change in the capacity of cells to respond to regulatory factors when entering the trunk formation stage. From a regulatory perspective, this might involve modification of transcription factor (TF) accessibility to their functional targets in the genome.

In this study, we aimed to understand the mechanisms involved in the switch from head to trunk development. For this, we compared transcriptome and chromatin accessibility profiles from the posterior epiblast region of wild type mouse embryos at embryonic day (E)7.5 and E8.5. We observed significant changes in transcriptomic profiles between these two stages. In addition to the expected changes in factors involved in pluripotency and in the *Hox* gene profiles, we observed modifications in a variety of functional groups, including the WNT signaling pathways, ubiquitination systems and lipid metabolic profiles that might interact together to change functional properties at the progenitor region of the embryo. Moreover, we uncovered a differential

requirement for Wnt palmitoleoylation upon trunk formation. We also observed major changes in chromatin accessibility profiles mostly involving intergenic regions, thus indicating a major switch in regulatory elements controlling head or trunk development, which were associated with changes in the binding activity of key transcription factors. In addition, we performed functional tests on specific enhancers identified in the chromatin analyses, including potential regulators of *Wnt5a* and *Pax3.* In transgenic reporter experiments these enhancers showed activity compatible with the regulation of the candidate target genes. However, when removed from the genome by edition procedures they had limited effect on the expression of those genes, indicating the existence of redundant enhancers that provide robustness to the system.

Results & Discussion

Transcriptome profile of the posterior epiblast in the developing embryo

To explore the changes in expression of genes involved in trunk formation, we used RNA-seq to obtain the transcriptome profiles from the posterior epiblast region of wild type mouse embryos at E7.5 and E8.5 (Fig 2.1A), representing respectively the progenitor-containing region of embryos before and after they engage in trunk formation. Principal component analysis separated the samples by timepoint (Fig 2.1B), revealing the presence of distinct transcriptomic profiles at these two developmental stages. Differential analysis revealed the presence of 2090 genes significantly downregulated, and 1668 genes upregulated at E8.5 relative to E7.5 (Fig 2.1C). Manual inspection of the list of differentially expressed genes (DEGs) identified downregulation at E8.5 of pluripotency genes, like *Pou5f1* or *Nanog* (Mitsui et al., 2003; Nichols et al., 1998), and genes involved in the initial establishment of the body axis and germ layers like *Tdgf1* (*Cripto*), *Nodal* or *Eomes* (Fig 2.1D) (Arnold et al., 2008; Conlon et al., 1991; Ding et al., 1998). Conversely, activation of central and posterior *Hox* genes was clearly observed at E8.5 (Fig 2.1E). These

findings fit with expression patterns reported for these genes, thus serving as an initial validation of our approach.

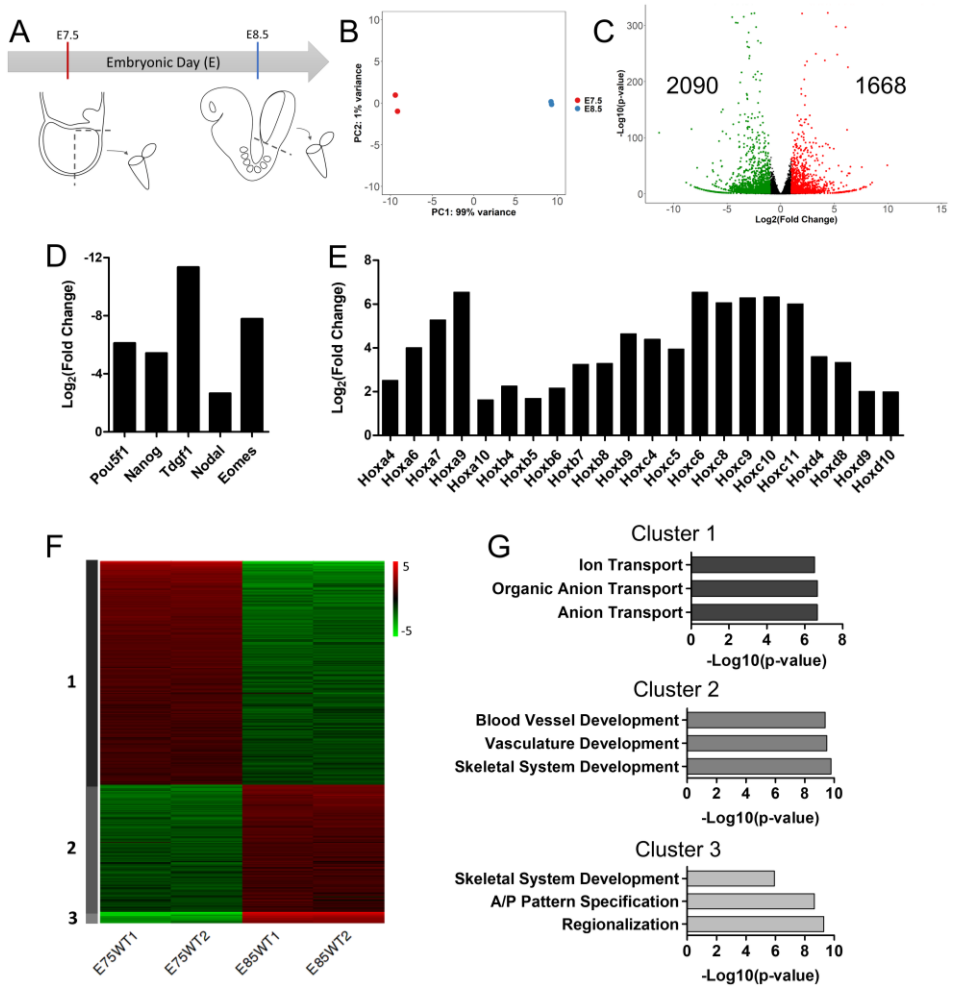


Figure 2.1 - Transcriptomic changes in the posterior epiblast associated with the head to trunk transition. (A) Schematic representation of sample collection from the posterior epiblast region of mouse embryos at E7.5 and E8.5. (B) Principal Component Analysis of RNA-seq data from E7.5 (red) and E8.5 (blue). (C) Volcano plot of RNA-seq gene expression ($|Log_2(Fold\ Change)| \geq 1$ & p-value < 0.05). Significantly upregulated genes at E8.5 are in red, downregulated at E8.5 are in green and non-significant in black. (D-E) Gene expression of key pluripotent and early developmental genes (D) and Hox genes (E). (F) K-means clustering of the 1000 most variable genes. Cluster 1: 616 genes; Cluster 2: 352 genes; Cluster 3: 32 genes. (G) Top 3 GO terms from biological processes associated with Cluster 1, 2 and 3.

K-means clustering of the top 1000 most variable genes produced three clusters with distinct gene expression dynamics (Fig 2.1F and G). Cluster 1 includes genes that became downregulated at E8.5; genes in this cluster are enriched in gene ontology (GO) terms related to anion and ion transport. Interestingly, a similar decrease in expression of genes enriched for ion transport and homeostasis has been described at the whole embryo level during the same stages analyzed here (Mitiku and Baker, 2007), further suggesting an important role for changes in ion transport profiles during early embryonic development. The full implication of this finding remains elusive. The control of ion fluxes has been implicated in patterning processes (Jaffe, 1981; Levin, 2021), including early stages in the establishment of left-right asymmetry associated with node activity (Levin et al., 2002; Raya et al., 2004). They also have been shown to control cell processes involved in cell migration, cell proliferation and autophagy (Becchetti et al., 2013; Schwab et al., 2012; Zhang et al., 2022). Focused experimental approaches will be required to explore if the drastic changes in ion transporter profiles observed in the progenitor-containing region during the head to trunk transition play a relevant role in the transition. Cluster 2 comprises genes moderately upregulated at E8.5, mostly associated with skeletal system, vasculature, and blood vessel development. Finally, cluster 3 is composed of genes strongly upregulated at E8.5. Genes in this cluster are enriched in skeletal system development, anterior/posterior pattern specification, and regionalization.

To get a closer image of the changes associated with the transition from head to trunk development, we built a protein-protein interaction (PPI) network (Fig 2.2A) based on the differentially expressed genes between E7.5 and E8.5, as obtained from StringDB (Szklarczyk et al., 2019). To focus our analysis on the most relevant interactions, we computed the metric backbone of this PPI network (Simas et al., 2021), which removed all redundant interactions and has been shown to help identifying genes and interactions responsible for core cellular programs (Correia et al., 2022a). Next, we identified structurally coherent network modules using LowEnDe (Correia et al., 2020), with an in-

house developed algorithm based on spectral decomposition and information theory. Our interpretation is that these network modules may represent core development functions that are responsible for key aspects of the head to trunk transition.

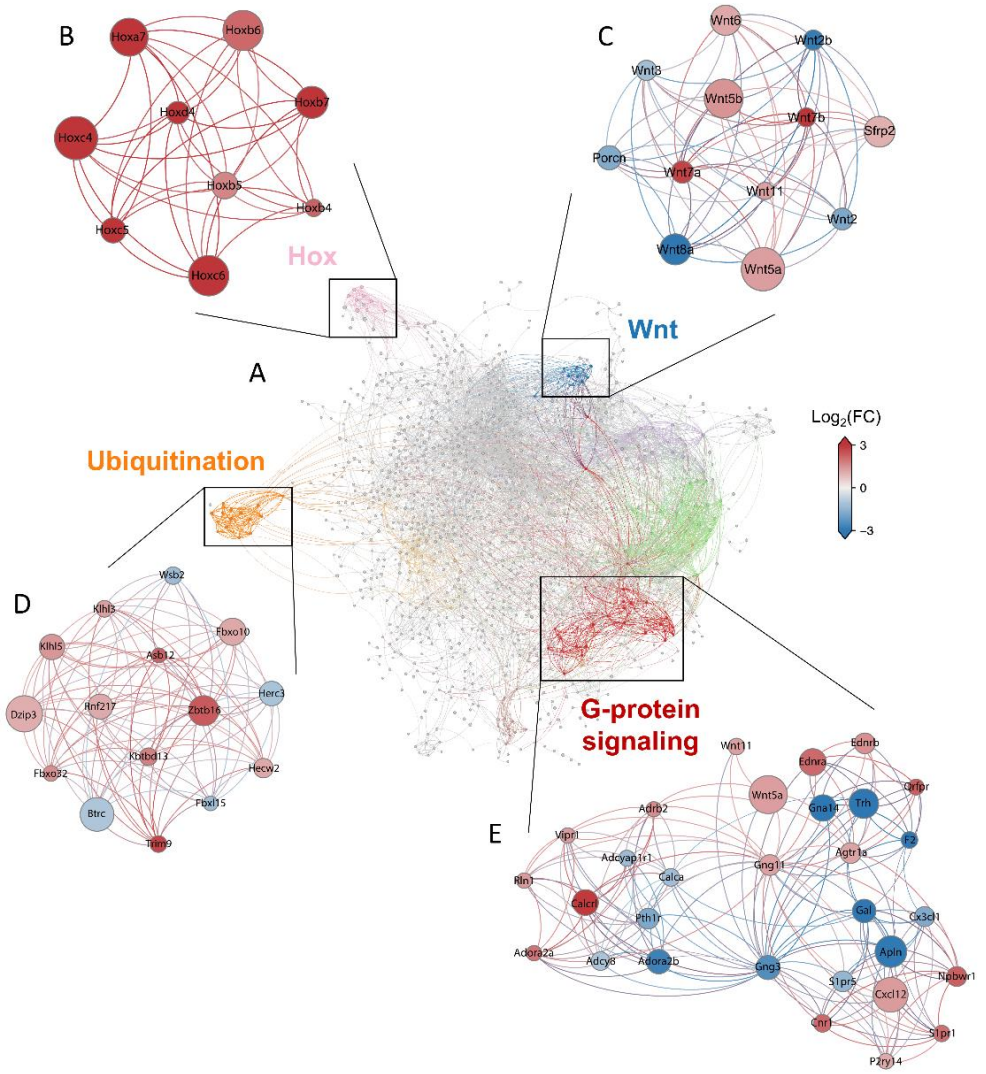


Figure 2.2 - Interaction networks reveal changes in various functional modules. (A) Protein-protein interaction network based on the differentially expressed genes between E7.5 and E8.5. Colored clusters represent structurally coherent network modules identified using LowEnDe (Correia et al., 2020). Purple cluster, Growth factors; Green cluster, Lipoprotein metabolism; Yellow cluster, Immune system. (B-D) Expanded versions of the Hox (B), Wnt (C), ubiquitination

(D) and G-protein coupled receptor signaling (E) modules are shown to highlight the genes included in each. Nodes are colored by Log_2(Fold Change), node size by Log_2(CPM). Significantly upregulated genes at E8.5 are in red, downregulated genes at E8.5 are in blue.

One of the resulting clusters comprised the *Hox* genes (Fig 2.2B) that we had already identified in our manual inspection of the differentially regulated genes, thus serving again as an internal validation of the approach. Another prominent cluster was associated with ubiquitination processes (Fig 2.2D) enriched in genes encoding for E3 ligases, the key determinants of substrate specificity of the ubiquitin proteasome system (Zheng and Shabek, 2017). This cluster contains a mix of up and downregulated genes, suggesting a switch in global ubiquitination patterns during the head to trunk transition that could impact general cellular functions by changing the availability of components involved in those processes. The large number of canonical WNT signaling components that have been shown to be regulated by ubiquitination, including β-catenin, Fzd, Lrp6, Dvl, Apc and Axin, (Abrami et al., 2008; Angers et al., 2006; Choi et al., 2004; Kim and Jho, 2010; Koo et al., 2012), suggests that changes in ubiquitination patterns might lead to functional modifications in WNT signaling during the head to trunk transition (see also below). Particularly interesting in this module is Btrc, known to promote β-catenin ubiquitination and its subsequent degradation (Hart et al., 1999; Kitagawa et al., 1999; Liu et al., 1999), which has been shown to also interact with components of several other signaling pathways and regulators of cell proliferation (Dorrello et al., 2006; Müerköster et al., 2005; Winston et al., 1999). Changes in *Btrc* levels might therefore affect interactions and activity balance between different signaling pathways, eventually impacting their global functional output. Indeed, the PPI network also identified several clusters composed of genes involved in different signaling pathways, indicating the existence of a substantial change in the signaling activities governing cell function when embryos engage in trunk formation.

One of those signaling-related clusters particularly prominent in the PPI network was composed of genes involved in G-protein coupled receptor signaling (Fig 2.2E). This cluster included both up- and down-regulated genes, suggesting a significant switch in G-protein-dependent signaling during the transition. Some changes involved ligands and receptors associated with G-protein-mediated signaling, like *Ednra, Ednrb, Adora2a, Calcrl* or *Agtr1a*, whose upregulation could be related to the development of the vascular system (Chigurupati et al., 2005; Liu et al., 2017; Schiffrin, 2007; Wong et al., 2003), or *Cxcl12,* which could be related to germ cell migration (Ara et al., 2003). Other changes in the G-protein-associated cluster could indicate a more general functional switch in the basic G-protein machinery. For instance, cluster analysis indicates a switch in the gamma subunits of the heterotrimeric G protein complexes, from Gng3 to Gng11, which could impact the selection of the pathways supported by the complex.

Those general changes in G protein-mediated signaling might play a role in the functional changes associated with WNT signaling during the head to trunk transition. In particular, the PPI network showed connections between the G-protein cluster and *Wnt5a* and *Wnt11*. This connection might expose a regulatory switch, considering that these Wnt factors are known to signal through non-canonical pathways (Abedini et al., 2020; Pandur et al., 2002; Slusarski et al., 1997) and their activity is essential when the embryo enters trunk development (Andre et al., 2015; Yamaguchi et al., 1999). It will be therefore interesting to determine whether the changes observed in the molecular composition of the G-protein signaling cluster from E7.5 to E8.5, promotes activation of the non-canonical WNT/Ca^{2+} pathway by Wnt5a and Wnt11 (Slusarski et al., 1997) when the embryo engages in axial extension. A more prominent involvement of the non-canonical WNT signaling downstream of Wnt5a when entering trunk development was also suggested by the upregulated *Sfrp2* expression at E8.5 (Fig 2.2C), since Sfrp2 redirects Wnt signals from Fz7 to Ror2, stabilizing the Wnt5a-Ror2 complexes that mediate Wnt5a activity during body axis development (Brinkmann et al., 2016; Ho et al.,

2012). The possible involvement of *Sfrp2* in this process is also supported by genetic data showing its requirement during trunk axial extension redundantly with *Sfrp1* (Satoh et al., 2006).

Another of the relevant changes in WNT signaling associated with the head to trunk transition is the switch from *Wnt*3 to *Wnt3a* functional dependency (Liu et al., 1999; Takada et al., 1994). In our datasets, *Wnt3* was downregulated at E8.5, fitting with its functional dynamics. *Wnt3a* expression levels, however, did not change from E7.5 to E8.5. This contrasts with the known *Wnt3a* functional requirements, as it is essential during trunk development but seems to be either inactive or functionally limited at earlier developmental stages given its inability to replace for *Wnt3* (Liu et al., 1999). This could suggest that stimulation of Wnt3a functional activity during axial extension might result from expression changes in additional factors modulating WNT signaling at different levels of the pathway. This is an intriguing possibility considering the differential effects that stabilization of Axin2 had on canonical WNT signaling in different areas of the embryo (Qian et al., 2011). In particular, it was shown that in embryos carrying an Axin2 allele (*canp*) coding for an Axin2 protein with increased stability, canonical WNT signaling was suppressed in the primitive streak during gastrulation and in the neural crest at later stages but was strongly up-regulated specifically in the progenitor zone of E8.5 embryos, eventually negatively impacting axial extension (Qian et al., 2011). These observations suggested fundamental changes in WNT signaling as embryos engage in axial extension, mostly involving the canonical pathway. A prominent candidate to be involved in differential Wnt regulation is *Porcn*, which codes for a molecule that introduces an essential palmitoleoyl moiety into a highly conserved serine residue of the Wnt ligands (Rios-Esteves et al., 2014; Takada et al., 2006).

Given the essential role of *Porcn* during gastrulation (Biechele et al., 2011), it was somewhat surprising to find a reduction of *Porcn* expression levels at E8.5, which was also observed in the expression patterns reported for this gene (Biechele et al., 2011). Whether this reduction plays a role in the WNT signaling switch associated with the head to trunk transition is unclear. Intriguingly, it has

been reported that pharmacological inhibition of Porcn impacted differently canonical and non-canonical WNT signaling in a cell line assay (Galli and Burrus, 2011), suggesting both that the Porcn-mediated modification might not be a universal requirement for WNT signaling and that *Porcn* expression levels might control the balance between canonical and non-canonical pathways.

WNT signaling dependency on Porcn during axial extension

We tested the effect of blocking Porcn activity on axial extension by incubating E8.5 embryos *in vitro* in the presence or absence of the Porcn inhibitor IWP-01. Our culture conditions allowed normal progression of development, with every control embryo attaining typical E9.5 morphology within 24 hours of incubation (Fig 2.3A). The presence of the inhibitor affected the development of every treated embryo. The brain structures were seriously reduced in size, likely affecting mainly the midbrain and anterior hindbrain structures, which also led to a substantial reduction in migratory cranial neural crest cells (Fig 2.3A and B). These features are consistent with the inhibition of Wnt1 signaling (Ikeya et al., 1997), thus serving as an internal control for IWP-01 activity. IWP-01 treated embryos underwent considerable extension at the caudal embryonic end, although they eventually became truncated. *Uncx4.1* expression indicated the presence of paraxial mesoderm along the whole anterior posterior axis, presenting fairly normal-looking somites for a considerable extent of the trunk, but losing segmental patterns towards the end of the axis (Fig 2.3J). The *Uncx4.1* signal almost reached the caudal embryonic end (Fig 2.3J), indicating that the presomitic mesoderm (PSM) was strongly reduced or absent, an idea also suggested by the lack of *Msgn1* signal (Fig 2.3D). *Sox2* expression indicated that IWP-01-treated embryos also developed a spinal cord, morphologically normal at the axial levels containing identifiable somites and becoming a wider flattened structure (Fig 2.3L-L'') in the region containing the disorganized *Uncx4.1* expression (Fig 2.3J). Importantly, even in the area showing abnormal neural and paraxial mesodermal patterns, IWP-01 treated embryos contained a single neural tube (Fig 2.3 J', L', N'). The axial

truncation in the context of a disorganized paraxial mesoderm and enlarged spinal cord could indicate an exhaustion of NMC cells derived from accelerated progenitor differentiation at the expense of self-renewal. The lack of *Cdx2* expression at the caudal end of IWP-01-treated embryos is consistent with this hypothesis (Fig 2.3F). Interestingly, *Shh* expression showed that the notochord also became truncated in the region where the paraxial mesoderm and the neural tube lose normal patterns (Fig 2.3N). *Tbxt* expression was reduced to a small spot beneath the neural tube (Fig 2.3H and H'), roughly corresponding to the position of the caudal end of *Shh* expression (Fig 2.3N and N'), indicating that it could represent the posterior end of the notochord.

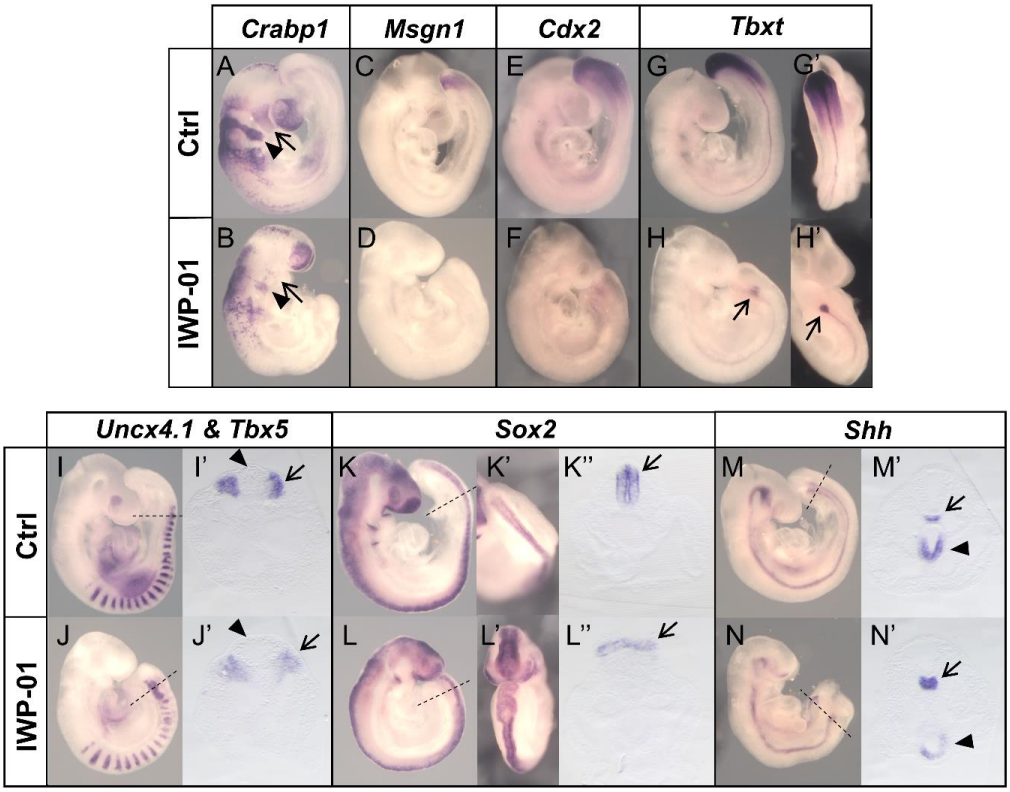


Figure 2.3 - Impact of Porcn activity during axial elongation. Embryos were cultured for 24hrs (E8.5 to E9.5) in the presence or absence of the Porcn inhibitor, IWP-01. Whole-mount *in situ* hybridization with *Crabp1* (A-B), *Msgn1* (C-D), *Cdx2* (E-F), *Tbxt* (G-H'), *Uncx4.1 & Tbx5* (I-J), *Sox2* (K-L) and *Shh* (M-N) probes. G', H', K' and L' show dorsal views of the posterior end of the respective embryos. I', J', K'', L'', M' and N' show transverse sections at the axial levels indicated

by the dashed lines in the respective embryos. Arrows and arrowheads in A and B indicate first and second branchial arches, respectively. *Tbxt* expression in the caudal region is reduced to a small spot (arrows in H and H'). Arrows and arrowheads in I' and J' indicate paraxial mesoderm and neural tube, respectively. The neural tube (arrows in K'' and L'') becomes a flat structure in the posterior end of the IWP-01-treated embryo. Arrows and arrowheads in M' and N' emphasize *Shh* expression in the notochord and gut, respectively.

Even considering that IWP-01 might not block Porcn activity completely, our data indicates that during axial extension WNT signaling involves a combination of Porcn dependent and independent activities. This contrasts with the essential role of Porcn during gastrulation (Biechele et al., 2011). Interestingly, the axial level at which the paraxial and spinal cord of IWP-01-treated embryos lose normal patterns roughly corresponds to the level when the embryo starts the transition into tail bud-dependent elongation, thus suggesting different requirements for the control of epiblast-driven and tail bud-dependent axial elongation. The differences between the reported phenotype of *Wnt3a* mutant embryos and the IWP-01 treated embryos indicate that during axial elongation Wnt3a signaling might include Porcn independent activities, an effect previously observed in a cell culture context (Rao et al., 2019). The presence of both neural and paraxial mesodermal tissues throughout the AP axis of IWP-01 treated embryos differs from the duplicated neural tubes replacing the paraxial mesoderm characteristic of the *Wnt3a* mutant embryos (Yoshikawa et al., 1997), thus indicating that the Wnt3a activity that modulates NMC cell fate might be Porcn-independent. Interestingly, IWP-01-treated embryos show a phenotype resembling *Wnt5a* mutants (Yamaguchi et al., 1999), indicating that its function during axial extension is Porcn-dependent. Together this suggests that the required equilibrium between differentiation and self-renewal of NMC cells might also entail proper balance of Porcn-dependent and Porcn independent WNT activities.

Chromatin accessibility landscape of the posterior epiblast in the developing embryo

To understand the regulation behind the changes observed in gene expression, we mapped global chromatin accessibility profiles. For this, we performed ATAC-seq (Corces et al., 2017) from tissues of the same regions and timepoints as those used for RNA-seq. Principal component analysis separated the samples by timepoint (Fig 2.4A), indicating the presence of distinct chromatin accessibility profiles at these two developmental stages. Both E7.5 and E8.5 datasets had a similar chromosomal distribution of accessible regions, with two thirds mapping to promoters and about 20% to intergenic regions (Fig 2.4B). Differential analysis of the two datasets identified 18197 regions with increased chromatin accessibility (open regions), and 11087 with decreased accessibility (closed regions) at E8.5 relative to E7.5 (Fig 2.4C). Interestingly, the differentially accessible peaks followed a distribution different to that observed for the individual datasets, with most peaks (57%) mapping to intergenic regions, 14% to introns and the contribution of promoters being reduced to around 21% (Fig 2.4B). This suggests that the transition between these developmental stages is to a large extent associated with a switch in regulatory elements. In addition, the finding that there are around ten times more genomic regions changing accessibility profiles than differentially expressed genes suggests a high complexity in the regulatory mechanisms controlling the transcriptional switch associated with the head to trunk transition.

From the regions showing differential accessibility, only 1418 could be associated with an annotated gene within 5 kb. Integrative analysis of transcriptomic and chromatin dynamics by crossmatching these 1418 regions with the differentially expressed genes (n=3758) identified 300 genes in common, of which 238 showed consistent regulation at both chromatin and transcriptomic levels (Fig 2.4D) (i.e., upregulated transcripts close to regions that became accessible or downregulated transcripts close to regions that lost accessibility). The remaining 62 regions might represent inhibitory elements.

These observations indicate that only a very small proportion of the regions that change accessibility during the head to trunk transition are predicted to control the closest annotated transcriptional unit, thus further complicating the understanding of the regulatory processes controlling the head to trunk transition.

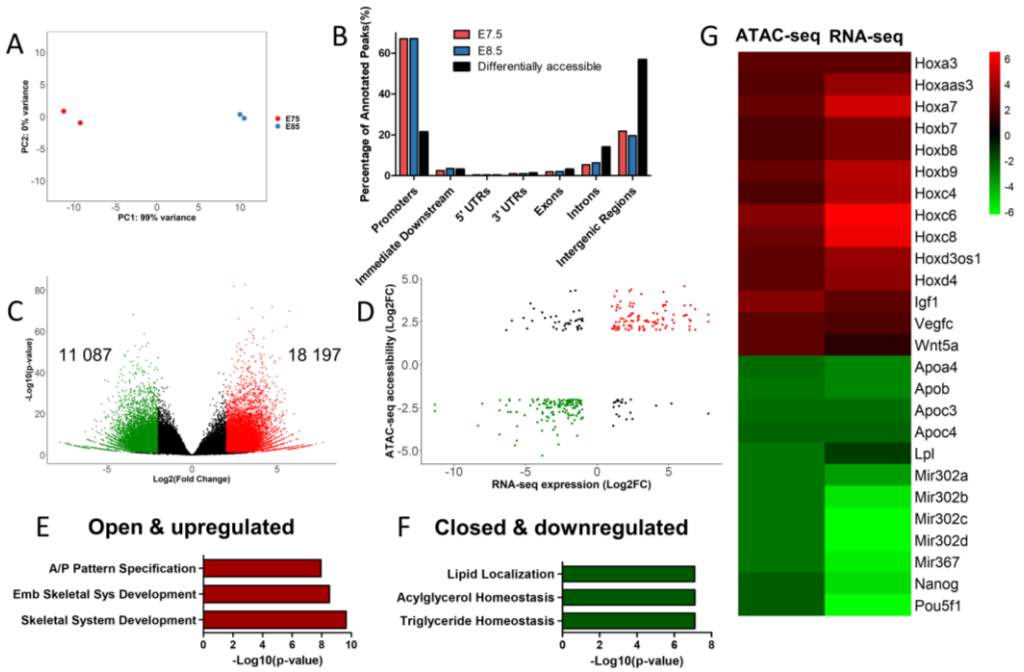


Figure 2.4 - Integration of genome accessibility and gene expression data. (A) Principal Component Analysis of ATAC-seq data from E7.5 (red) and E8.5 (blue). (B) Genomic distribution of ATAC-seq peaks identified at E7.5, E8.5 and distribution of only the differentially accessible peaks (black). (C) Volcano plot of ATAC-seq peaks (|Log2(Fold Change)|≥2 & p-value < 0.05). Significantly open regions at E8.5 in red, closed regions at E8.5 in green and non-significant in black. (D) Scatterplot showing correlation between genomic accessibility and gene expression. Significantly accessible and upregulated genes in red, closed and downregulated genes in green. (E) Top 3 GO biological process terms of positively regulated genes at E8.5 (red group in D), Abbreviations: Emb, Embryonic; Sys, System, A/P, Anterior/Posterior. (F) Top 3 GO biological process terms of negatively regulated genes at E8.5 (green group in D) (G) Heatmap of Log2(Fold Change) of ATAC-seq and RNA-seq levels.

Analysis of GO terms of this restricted group revealed an enrichment in anterior/posterior pattern specification and skeletal system development, in genes which are both accessible and upregulated at E8.5 (Fig 2.4E). These include several *Hox* genes, most particularly those of central and posterior paralog groups (Fig 2.4G), which might reflect the activation of enhancers within the *Hox* clusters upon sequential global opening of the clusters during axial extension (Beccari et al., 2016). The group of less accessible and downregulated genes include genes related to stem cell pluripotency and proliferation (Fig 2.4G), like the already mentioned *Pou5f1* and *Nanog.* This is consistent with the known position of relevant regulatory regions for these genes (Agrawal et al., 2021; Okazawa et al., 1991). This group also included the *miR-302/367* cluster, important for stem cell maintenance and repression of cell differentiation (Rosa and Brivanlou, 2011).

GO terms of the less accessible and downregulated genes were enriched for triglyceride homeostasis and lipid metabolism (Fig 2.4F), including several *Apo* genes as well as *Lpl*, that catalyzes the hydrolysis of triglycerides (Fig 2.4G). These observations indicate that the head to trunk transition is associated with changes in lipid metabolism, which have the potential to impact the activity of various signaling pathways. For instance, lipid modifications have been shown to be essential to generate functionally competent Wnt and Hedgehog molecules (Gallet et al., 2003; Lu et al., 2018). In the case of Wnt ligands, they contain several lipidic modifications, including the above-mentioned palmitoleoylation, which have been shown to affect differently the functional activity of different Wnt molecules (Doubravska et al., 2011; Galli and Burrus, 2011) and, as already discussed above, could be involved in the implementation of the functional switch in WNT signaling associated with the head to trunk transition. Interestingly, in Drosophila embryos lipid-modified Hedgehog and Wingless require association with lipoproteins for long-range spreading of their activity (Panáková et al., 2005), and Wnt5a has also been shown to associate with lipoprotein particles for long distance regulation of hindbrain development (Kaiser et al., 2019). In our datasets, several genes

encoding for lipoprotein components are downregulated at E8.5. While this could be related to changes in the transport of lipid nutrients to the developing embryo, as shown for Apob during mouse embryogenesis (Farese et al., 1995), it could also impact Wnt and Hedgehog activities by determining the spatial range of their activity at different developmental stages. In this regard, it is interesting to note that one of the first determinants of left-right asymmetry involves a flow in the node that has been shown to include lipoprotein vesicles containing Shh (Tanaka et al., 2005). This occurs around E7.5, although signs of asymmetry are only apparent later in development (Okada et al., 2005; Zhang et al., 2001). Therefore, the reduction in apolipoprotein-encoding genes could be involved in a restriction of the timing of this signaling. The Wnt-coreceptors Lrp5 and Lrp6 also belong to the family of lipoprotein receptors, thus adding the potential involvement of apolipoproteins in the differential regulation of WNT signaling by modulating interactions between the Wnt molecules and their receptors.

Transcription factor binding activity in the posterior epiblast

To assess how the modification of the chromatin accessibility profiles between E7.5 and E8.5 was reflected in the binding profiles of TFs known to be involved in developmental processes, we searched for TF footprints in our ATAC-seq datasets using HINT-ATAC (Li et al., 2019). We found several TFs with a significant difference in activity score between the two developmental stages (Fig 2.5A). At E7.5 we observed a higher activity score for TFs involved in pluripotency, like Pou5f1, Nanog and Sox2. The average ATAC-seq profiles around the binding sites of each of these TFs revealed that, although at a lower level and in a reduced number of regions, binding activity was still detected at E8.5 (Fig 2.5B-D). This might reflect a change in the functional profile of those factors as development proceeds. For instance, while Sox2 and Pou5f1 are required for pluripotency (Avilion et al., 2003; Nichols et al., 1998), later in development they are involved in trunk elongation (Pou5f1) or in neural tube development (Sox2) (Aires et al., 2016; DeVeale et al., 2013; Graham et al.,

2003). At E8.5 the highest activity scores were provided by Cdx2, Cdx1, and several posterior Hox proteins (Fig 2.5A). Interestingly, their ATAC-seq profiles showed shallow footprints at E7.5 (Fig 2.5E-G), revealing that binding of these factors to their genomic targets mostly starts when the embryo engages in trunk development. These observations fit the genetic data showing that in the absence of Cdx activity, mouse embryos are truncated at the head to trunk transition (Amin et al., 2016; Chawengsaksophak et al., 1997; Savory et al., 2011), thus indicating that the functional requirement for these genes starts at this transition. Conversely, the binding profile of Tbxt (Brachyury), another of the main regulators of axial extension (Herrmann et al., 1990), was similar at E7.5 and E8.5 (Fig 2.5H). This suggests that Tbxt transcriptional activity might be similar at both developmental stages despite only being essential when the embryos enter trunk development (Amin et al., 2016; Chesley, 1935). Alternatively, the Tbxt binding activity identified with the footprints at E7.5 might not be directly related to this protein but to Eomes, a T-box TF essential for endodermal and mesodermal development at early developmental stages (Arnold et al., 2008) that shares DNA target sequence with Tbxt. Discerning between these possibilities will require evaluating Tbxt binding profiles in *Eomes* mutant embryos. Nfya and Nfyb were also found to have a higher activity score at E8.5 (Fig 2.5I). *Nfya* homozygous mutants exhibit early embryonic lethality (Bhattacharya et al., 2003), possibly related to the role of Nfya in zygotic genome activation (Lu et al., 2016). Besides the role in early development, our data suggests that *Nfya* and *Nfyb* may also be important at the transition stage. Together, these results highlight a change in the main regulatory networks involved in each of these developmental stages, which is reflected by the activity levels of specific TFs.

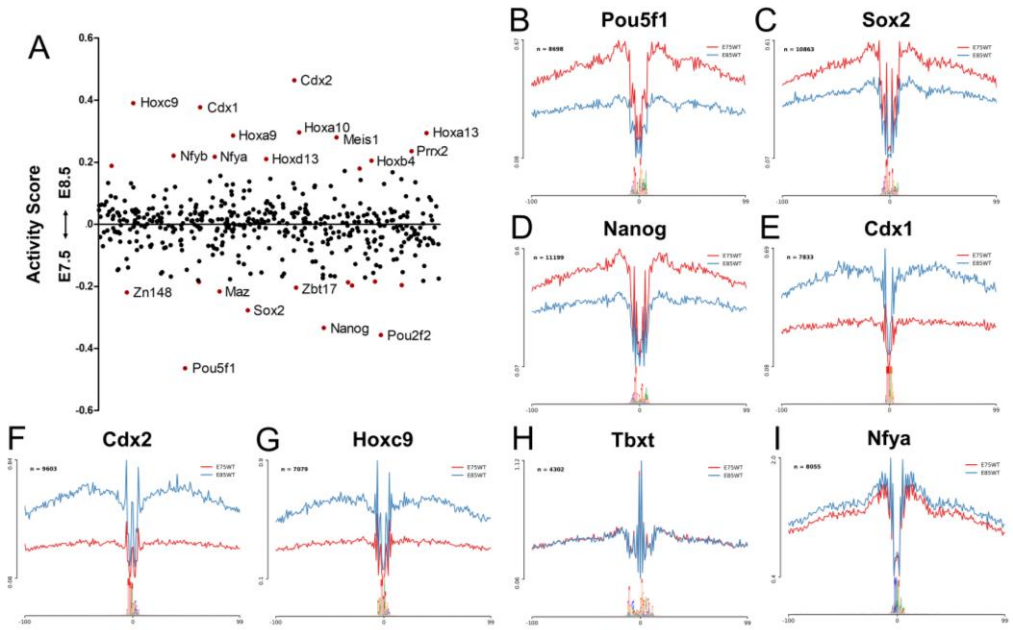


Figure 2.5 - TF activity dynamics during the head to trunk transition. (A) Scatter plot of TF activity dynamics between E7.5 and E8.5. The y-axis represents the differences in TF binding activity. Each point represents a TF, points colored in red have significantly different activity scores (p-value < 0.05). Labelled points have a differential |Activity Score| > 0.2. (B-I) Average ATAC-seq profiles of Pou5f1 (B), Sox2 (C), Nanog (D), Cdx1 (E), Cdx2 (F), Hoxc9 (G), Tbxt (H) and Nfya (I) binding sites. Red profiles correspond to E7.5, blue profiles to E8.5, n indicates the number of binding sites used to calculate the average profiles.

Testing potential enhancer regions

From the differentially accessible peaks identified between E7.5 and E8.5 we assembled a list of phylogenetically conserved regions (Pollard et al., 2010), that mapped close to important developmental genes, and evaluated their capacity to promote activity patterns consistent with a role in the head to trunk transition. For this, we cloned the genomic sequence of each candidate region into a plasmid with a minimal promoter and the *β-galactosidase* reporter gene and examined their activity in transgenic mouse embryos. Of the 10 regions tested, six elements induced consistent expression of the reporter gene in transgenic embryos (Table 2.1). Peak_92869 promoted reporter expression in the forebrain and optic vesicles (Fig 2.6A). Peak_123246 activated expression

in the lateral mesoderm, limited to the trunk region (Fig 2.6B). Peak_57185 and Peak_57188 showed similar activity patterns, in the posterior epiblast, neural tube, lateral mesoderm, PSM and somites, although expression in the latter tissue extends more anteriorly in *Peak_57185-β-gal* (Fig 2.6C and D). These regions are located in an intron and downstream of *Hoxc9*, respectively, and their pattern of reporter expression fits with a possible role in regulating *Hoxc9* expression.

Table 2.1 – List of genomic regions tested in *β-galactosidase* reporter assays.

ID	Chr	Start	End	Regulation at E8.5	Reporter expression
Peak_37	chr1	4 506 473	4 507 113	Closed	No
Peak_621	chr1	16 127 242	16 127 499	Open	No
Peak_3705	chr1	78 171 700	78 172 183	Open	Yes
Peak_45848	chr14	28 500 913	28 502 135	Open	Yes
Peak_57185	chr15	102 982 949	102 984 110	Open	Yes
Peak_57188	chr15	102 986 211	102 988 795	Open	Yes
Peak_83746	chr2	74 685 149	74 685 691	Open	No
Peak_92869	chr3	34 645 894	34 646 882	Closed	Yes
Peak_102158	chr4	47 449 347	47 450 598	Closed	No
Peak_123246	chr6	52 173 674	52 174 119	Open	Yes

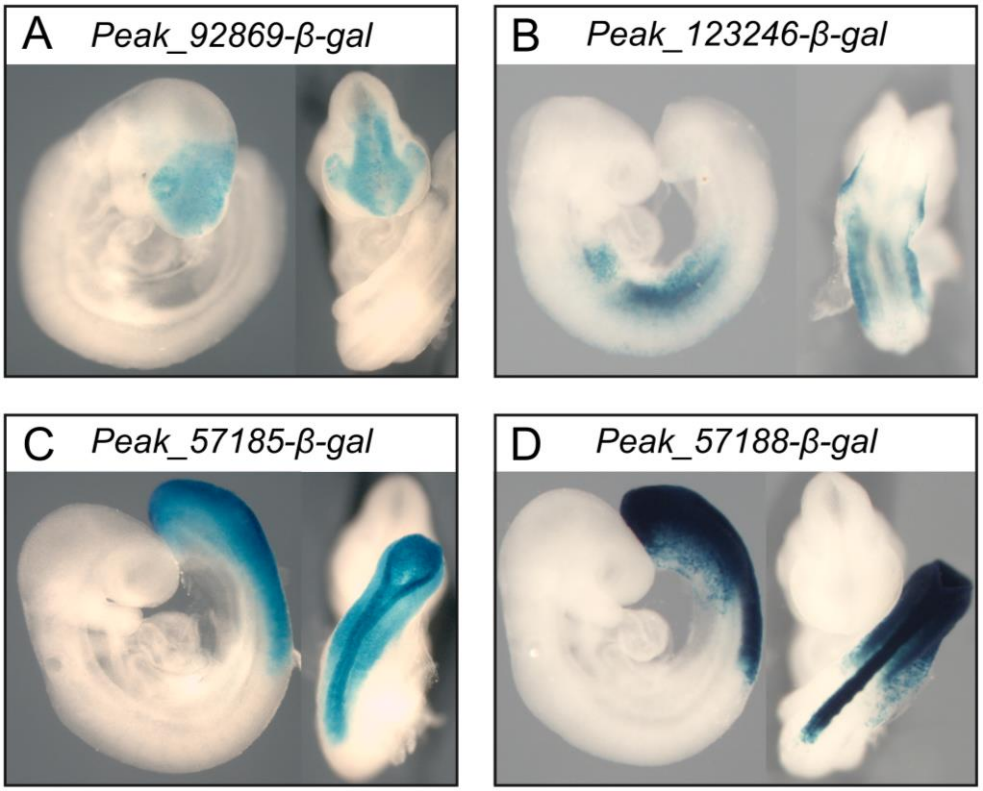


Figure 2.6 – *β-galactosidase* reporter activity in transgenic embryos. Transgenic embryos for reporter constructs with Peak_92869 (A); Peak_123246 (B); Peak_57185 (C) and Peak_57188 (D).

Transgenic embryos featuring Peak_45848 consistently induced reporter expression in the posterior epiblast and emerging neural tube, in a pattern closely resembling *Wnt5a* expression (Fig 2.7A). We will refer to this region as CR1. CR1 is approximately 3.3kb upstream of *Wnt5a* transcriptional start site and becomes accessible at E8.5 (Fig 2.7B). In addition, *Wnt5a* expression is upregulated at E8.5 (Fig 2.2C) which would be consistent with involvement of CR1 in *Wnt5a* activation in the progenitor region during the head to trunk transition.

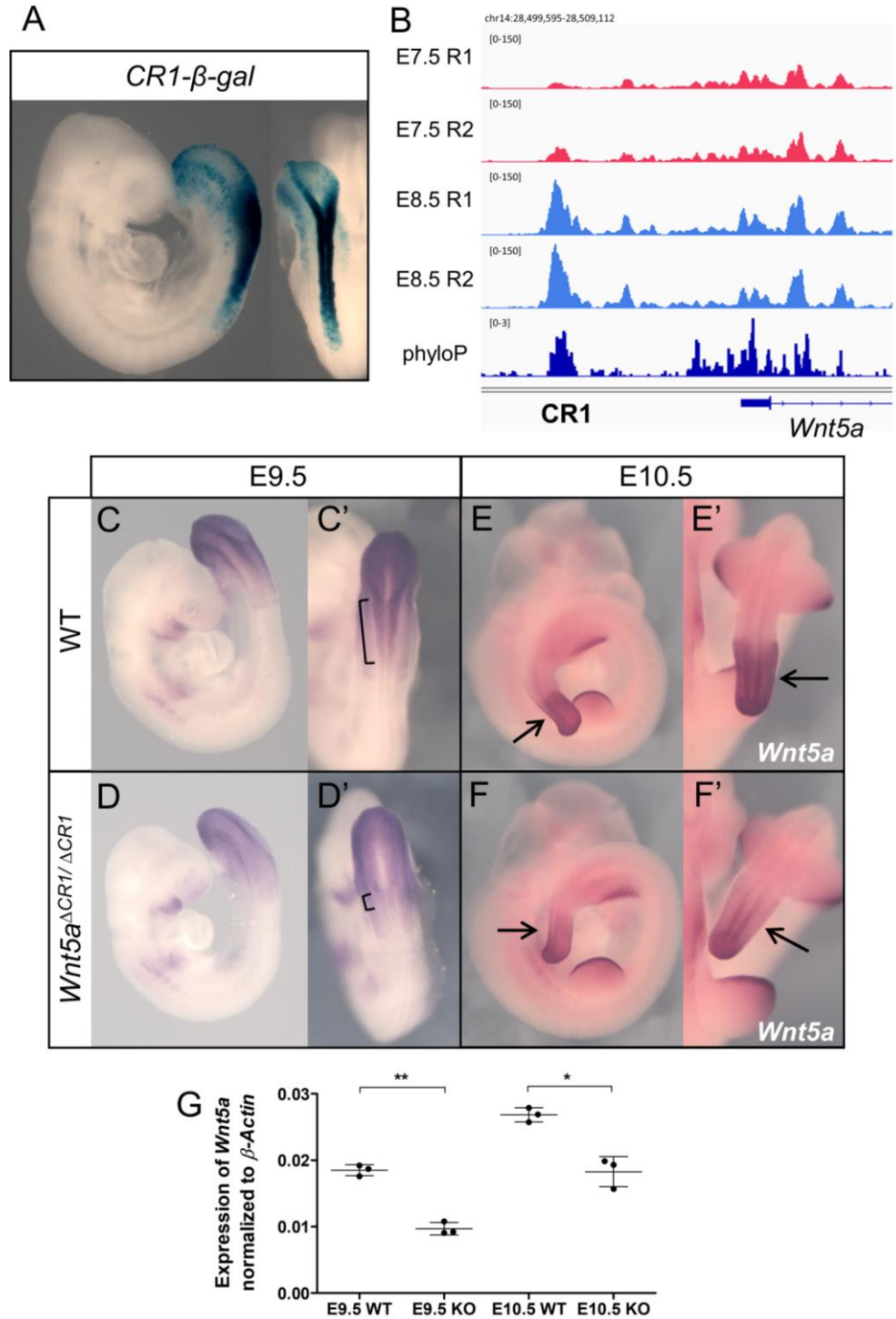


Figure 2.7 - Characterization of Wnt5a enhancer, CR1. (A) β-gal staining of *CR1-β-gal* transgenic embryo, also showing dorsal view of the caudal region. (B) ATAC-seq tracks showing

accessibility profiles in the CR1 region. Phylogenetic conservation data (phyloP) (Pollard et al., 2010) is shown in dark blue. (C-F') Whole-mount in situ hybridization of wild type and *Wnt5a*$^{\Delta CR1/\Delta CR1}$ embryos at E9.5 and E10.5 using a probe for *Wnt5a*. (C' and D') Dorsal view of the posterior end of the embryo emphasizing the reduction of *Wnt5a* expression in the neural tube (brackets). At E10.5 *Wnt5a* expression is reduced in the PSM (arrows in E' and F'). (G) RT-qPCR analysis of *Wnt5a* gene expression in wild type (WT) and *Wnt5a*$^{\Delta CR1/\Delta CR1}$ (KO) embryos at E9.5 and E10.5. *Wnt5a* expression is normalized to *β-Actin*. Error bars indicate the standard deviation; **, p-value < 0.01; *, p-value < 0.05

To directly explore this hypothesis, we generated CR1 deletion mutants (*Wnt5a*$^{\Delta CR1}$). Whole-mount *in situ* hybridization suggested a reduction in *Wnt5a* expression levels in the neural tube of *Wnt5a*$^{\Delta CR1/\Delta CR1}$ embryos at E9.5 (Fig 2.7C' and D'), and in the PSM at E10.5 (Fig 2.7E-F'). This downregulation was confirmed by quantitative RT-PCR, at both stages (Fig 2.7G). These results suggest that, while the CR1 element participates in the regulation of *Wnt5a* expression *in vivo*, this regulation should also involve the activity of additional redundant enhancers that confer robustness to *Wnt5a* expression, able to keep a baseline *Wnt5a* transcription in *Wnt5a*$^{\Delta CR1/\Delta CR1}$ mutants, thus allowing their full embryonic development. Despite the observed downregulation of *Wnt5a*, *Wnt5a*$^{\Delta CR1/\Delta CR1}$ mutants developed normally, generating adult animals with no obvious phenotypic defects. This contrasts with *Wnt5a*$^{-/-}$ mutants, where loss of *Wnt5a* leads to perinatal lethality, with embryos showing an absence of tail and a shortening of the anterior-posterior axis (Yamaguchi et al., 1999).

Finally, Peak_3705 (We will refer to it as CR2) showed reporter expression in the neural tube and midbrain, in a pattern similar to *Pax3* expression in these tissues (Fig 2.8A). CR2 is located in intron 4 of *Pax3* and becomes accessible at E8.5 (Fig 2.8B). Since *Pax3* is the only annotated gene within a highly interactive TAD sub-domain, CR2 is most likely regulating the expression of *Pax3* itself (Fig 2.8C), which we found upregulated at E8.5. To evaluate CR2 function *in vivo*, we generated CR2 deletion mutants (*Pax3*$^{\Delta CR2}$). However, whole-mount *in situ* hybridization revealed no changes in *Pax3* expression patterns when compared with wild type embryos (Fig 2.8D). In addition,

embryonic development of homozygous mutants was unaffected and adult animals showed no obvious phenotype, indicating that these mutants can still maintain *Pax3* expression through redundant enhancers. Similar reporter assays have previously identified enhancer regions upstream of *Pax3,* directing expression in neural crest precursors and interlimb somites, however the impact of deleting these elements *in vivo* was not investigated (Brown et al., 2005; Milewski et al., 2004).

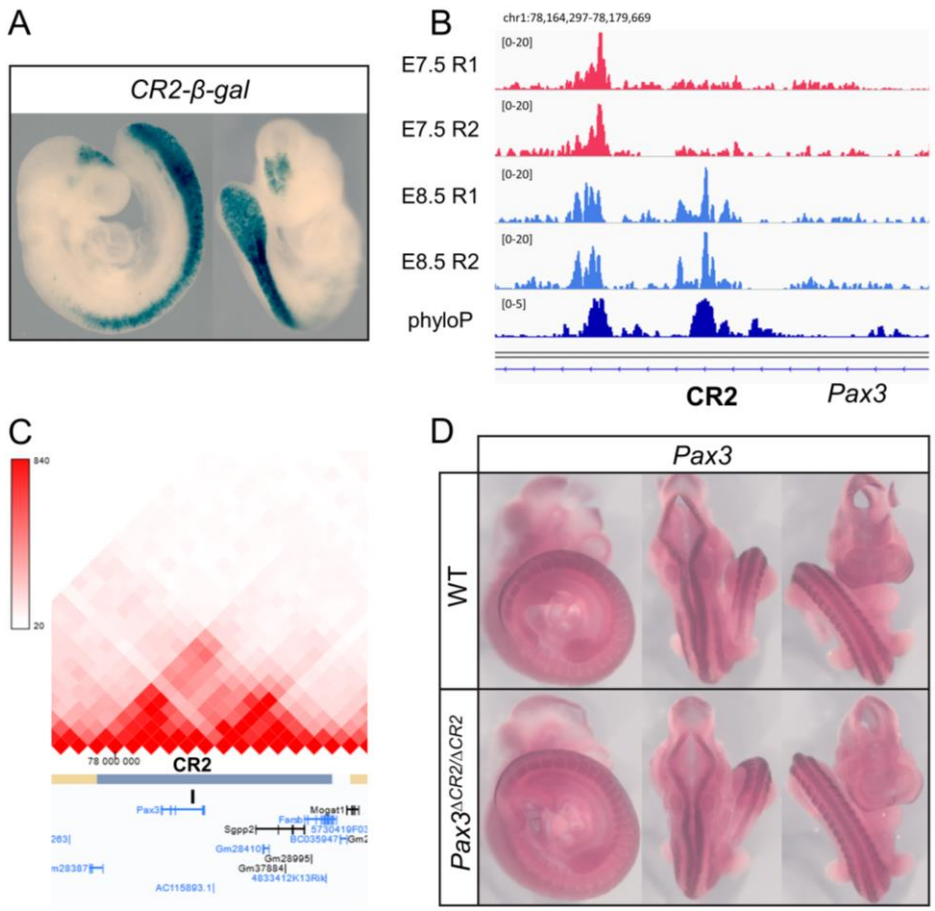


Figure 2.8 - Characterization of CR2. (A) β-gal staining of *CR2-β-gal* transgenic embryo, also showing dorsal view of the caudal region. (B) ATAC-seq tracks showing accessibility profiles in the CR2 region. Phylogenetic conservation data (phyloP) (Pollard et al., 2010) is shown in dark blue. (C) Hi-C data from 3D genome browser (Wang et al., 2018) demonstrating CR2 location in the same TAD as *Pax3*. (D) Whole-mount *in situ* hybridization of wild type and *Pax3*$^{\Delta CR2/\Delta CR2}$ embryos with a probe for *Pax3*.

Conclusion

Overall, this study provides a comprehensive analysis to gain insights into the mechanisms regulating the remarkable changes in tissue activity associated with the transition from head to trunk development, combining differential screening with bioinformatic treatment of the resulting data. The specific treatment of the DEG profiles with the identification of modules in the PPI network revealed changes in WNT signaling, ubiquitination and the basic machinery of G-protein-mediated signal transduction that could engage in interactions resulting in a global functional output. We also observed a switch in the functional relevance of Wnt palmitoleoylation by Porcn, essential during gastrulation (Biechele et al., 2011; Takada et al., 2006) but later becoming differentially required for the control of axial extension and progenitor differentiation during trunk formation.

The high disparity in the number of changes in chromosomal accessibility and differentially expressed genes at the two developmental stages indicates that the control of the changes in gene expression might also be very complex. Consistent with this, the lack of obvious phenotypes upon deletion of regulatory elements that largely reproduce the expression profiles of the genes they might regulate, argues for the existence of a considerable degree of redundancies among regulatory mechanisms. This redundancy, which has been previously observed for other regulatory regions (Cunningham et al., 2018; Osterwalder et al., 2018), might confer robustness to developmental processes by providing protection against genetic and environmental perturbations (Antosova et al., 2016; Frankel et al., 2010; Osterwalder et al., 2018; Perry et al., 2010). Future studies testing the effects of CR1 or CR2 deletion combined with a heterozygous inactivation of the *Wnt5a* or *Pax3* genes, respectively, or with other potential regulatory elements may reveal phenotypic traits normally suppressed by functional redundancy among enhancers. In addition, our datasets can be used as a resource for future research not only to explore the role of other enhancer regions but also to delve into the mechanisms of gene

regulatory networks involved in the head to trunk transition by combining it with gene knockout studies.

Methods

Mice and Embryos

Mouse embryos were recovered by cesarean section at different developmental stages and processed accordingly for the distinct analyses described below. All experiments conducted on animals followed the Portuguese (Portaria 1005/92) and European (Directive 2010/63/EU) legislations concerning housing, husbandry, and welfare. The project was reviewed and approved by the Ethics Committee of *Instituto Gulbenkian de Ciência* and by the Portuguese National Entity, *Direcção Geral de Alimentação e Veterinária* (license reference 014308).

CR1 and CR2 deletion mice were generated by CRISPR/Cas9 (Wang et al., 2013) on the FVB/J background as previously described (Tekko et al., 2022), using two gRNAs targeting the border of the sequence to be deleted and one ssDNA oligo bridging the two sides of the deletion (Table 2.2) to increase the edition efficiency. In these cases, each gRNA was generated by annealing the relevant Alt-R-CRISPR-Cas9 crRNA (targeting sequences in Table 2.2) with the Alt-R-CRISPR-Cas9 tracrRNA (all purchased from IDT). 1μM of each gRNA was incubated with 100 ng/μl of theCas9 protein and 10 ng/μl of the replacement DNA and microinjected into the pronucleus of fertilized mouse oocytes. Identification of deletion mutants was performed by PCR using the oligonucleotides specified in Table 2.3. Positive founders were crossed with wild type mice to generate F1 heterozygous animals that were then used to build the mutant lines. Homozygous mutants were then generated by heterozygous crosses. Mice and embryos were genotyped by PCR using primers specified on Table 2.3.

Table 2.2 - gRNA and ssDNA used for CRISPR/Cas9.

CR1	gRNA_CR1_1	CCAGTGGCAGTATTCTGTGA
	gRNA_CR1_2	CTGTGTAGCCGTAGTTTGCC
	ssDNA	CCCCCTAACCTCAAGGGAGCCTTTGTCCCCCACAGGCTAGTGGC CAGTGGCAGTATTCTGGCCAGGAGGTGAGGGACTTCCACAAACT GGAGGCTCTCCTTTGGGAGTCTTCCCCAGTGG
CR2	gRNA_CR2_1	CACTTCATCCAAGCACGAGA
	gRNA_CR2_2	GCTCGAGAGAGAGGAAGGAG
	ssDNA	TGCTCCTTTTTAAATGTAAGGCTTCCTCCTCAAACCTTACCATCAC TTCATCCAAGCACGCTTCCTCTCTCTCGAGCCAGCCTGGTGGGA CGTGGGAGCAGAATTAATAAACAAAATAAG

Table 2.3 - Primers used for genotyping.

CR1	KO	Fw	GTCTCTTCCATGAGTGCTGAG
		Rv	CTGCATTCTAAGAAGCAGTCC
	WT	Fw	ACCCACTTTCTACAGCAGATC
		Rv	CTGCATTCTAAGAAGCAGTCC
CR2	KO	Fw	CTAGCAGACAACACTGACCTG
		Rv	ACGAACTTACGTACTCTGGTG
	WT	Fw	CTAGCAGACAACACTGACCTG
		Rv	CATCCTGCATTTAGGCAATGC

RNA-sequencing analysis

Posterior epiblasts of wild type mouse embryos at E7.5 and E8.5 were dissected and snap frozen. Total RNA was isolated from pooled samples with TRI Reagent following the manufacturer's protocol. RNA samples were then resuspended in RNase-free water. RNA concentration and purity were determined on an AATI Fragment Analyzer (Agilent). RNA-seq from E7.5 and E8.5 tissues was performed using two separate biological replicates. Libraries were prepared from total RNA using the SMART-Seq2 protocol (Picelli et al., 2014). Sequencing was performed on Illumina NextSeq500, generating >25M single-end 75 base reads per sample. Reads were aligned to the reference mouse genome (mm10) using STAR (Dobin et al., 2013). Read count normalization and differential expression between samples was analyzed using

DESeq2 (Love et al., 2014). RNA-seq data is available in the GEO accession database under the accession number GSE220246. K-means clustering was performed on the 1000 most variable genes using the standard R function 'kmeans()'. The elbow method was used to determine the number of clusters to use for this analysis. Gene ontology enrichment analysis was performed using PANTHER (Ashburner et al., 2000; Mi et al., 2019), by selecting for biological processes using Fisher's Exact test and False Discovery Rate. No background gene list was used. Gene ontology results presented are ranked by Fold Enrichment.

To assemble the PPI network, the DEG were filtered according to the following criteria: log of count per million (logCPM) >1; absolute log fold-change (logFC) >1; and false discovery rate (FDR) <0.05. All possible interactions between DEGs were retrieved from the STRING v11 protein-protein interactions database (Szklarczyk et al., 2019). The mouse transcriptome network was then constructed from the set of expressed genes and their corresponding STRING PPI. We casted this network as a weighted graph, where edge weights (given by the STRING PPI scores) denote the probability of the connected genes interacting and thus jointly contributing to a specific function. To remove redundant edges and focus our attention on the most important interactions we extracted the (metric) backbone of the mouse transcriptome network (Simas et al., 2021). The metric backbone is a subgraph that is sufficient to compute all shortest paths in the network, thus removing edges that break the triangle unequally (and are therefore redundant regarding the shortest paths). This network retains all metric edges and preserves all the nodes in the original network (Correia et al., 2022b; Simas et al., 2021). This metric backbone of transcriptome networks has been previously used to successfully identify biologically relevant genes and their interactions (Correia et al., 2022a). Network modules, i.e., structurally coherent structures in the transcriptome network backbone were identified using LowEnDe (Correia et al., 2020), an algorithm based on the spectral decomposition of the adjacency matrix coupled with information theory to identify overlapping modules in

weighted graphs. Importantly, in this method genes may participate in more than one module at the same time, reflecting the possible participation of genes in multiple cell functions.

Embryo culture with Porcn inhibitor

Wild type E8.5 embryos were dissected, in cold GMEM (Sigma, G5154), keeping the yolk sac intact. Embryos were cultured in 60% Rat serum, 40% GMEM and Pen/Strep (Gibco, 15070063). For embryos cultured with Porcn inhibitor, 500 nM of IWP-01 (MedChem express, HY-100853) was added as in (Galli and Burrus, 2011), whereas for control embryos an equal volume of DMSO was added. Embryos were cultured for 24 hours in a rotator bottle culture apparatus (B.T.C. Engineering, Milton, Cambridge, UK) at 37˚C, 65% O_2. Three embryos were cultured per tube in 1,5 ml of media. Embryos were collected after 24hrs, dissected and fixed in 4% PFA at 4˚C overnight. They were then processed for *in situ* hybridization, 2 embryos were stained per probe and condition, showing similar patterns. In addition, the structure of the neural tube was also assessed in the sections of embryos stained for other markers, showing highly reproducible patterns.

ATAC-seq

Posterior epiblasts of mouse embryos were collected to 500 μl of cold M2 (Sigma, M7167), spun down to remove supernatant and incubated with 500 μl of Accutase (Sigma, A6964) for 30 min at 37ºC, with shaking at 600 rpm, to dissociate the tissue into single cells. ATAC-seq was performed as previously described (Corces et al., 2017), using two separate biological replicates for each condition. The amplified libraries were double-step size selected (0.5x followed by 1x) using SPRIselect (Beckman Coulter, B23317) according to manufacturer's instructions. Pooled ATAC-seq libraries were sequenced on a NextSeq500 (Illumina) at 50M paired-end 75 base reads per sample.

ATAC-seq Data Analysis

Fastq files were processed with GUAVA v1, following the recommended guidelines (Divate and Cheung, 2018). GUAVA enables pre-processing of raw sequencing reads, mapping of reads to a reference genome, peak calling and annotation, as well as differential analysis between samples. ATAC-seq data is available in the GEO accession database under accession number GSE220245. All genome browser tracks were captured using Integrative Genomics Viewer (Robinson et al., 2011). Phylogenetic conservation data for multiple alignments of 59 vertebrate genomes to the mouse genome (mm10.60way.phyloP60way) was obtained from phyloP directory (Pollard et al., 2010) and loaded into IGV. Hi-C data was obtained from 3D genome browser (Wang et al., 2018) using the 'mm10 mESC Bonev_2017-raw' dataset at 40kb resolution. To visualize our candidate regions within the context of this dataset we loaded a BED file with the coordinates of our candidate regions to UCSC Genome Browser (Kent et al., 2002) and loaded this session into the 3D genome browser.

ATAC-seq data was analyzed for TF footprints using HINT (Li et al., 2019). Replicates were merged to increase read depth and processed with "rgt-hint footprinting" command. Footprint motifs were matched to HOCOMOCO database (Kulakovskiy et al., 2018) with "rgt-motifanalysis matching" and then further assessed for differential motif occupancy with the "rgt-hint differential" command.

β-Galactosidase Transgenics

For reporter analyses, candidate regions identified by ATAC-seq data were amplified by PCR from mouse genomic DNA (primers provided below, Table 2.4) and cloned upstream of a cassette containing the adenovirus 2 minimal late promoter, the β-galactosidase cDNA, and the polyadenylation signal from SV40 (Jurberg et al., 2013). Transgenic mice were produced by pronuclear injection (Hogan et al., 1994). The β-galactosidase staining was performed as previously described (Jurberg et al., 2013). At least 3 transgenic embryos

displaying consistent reporter expression patterns were observed for each candidate.

Table 2.4 - Primers used to amplify candidate regions for *β-Galactosidase* assays.

CR1	Fw	TACTCGAGCTGCTGCTCTTGACTCTGAAG
	Rv	ATCTGCAGATGCTCTGGACTCCGAGGAAC
CR2	Fw	ACCTCGAGATGTCAACGAAAGAAAGACGC
	Rv	TTCTGCAGGAGGTGTAGCTTTAAGCACAC
Peak_37	Fw	ACCTCGAGTCAAAGGGTAACTAAGAACGC
	Rv	CGAGATCTTTGCTAGTGGGTATAGGAATC
Peak_621	Fw	TACTCGAGGACCAGCATTGGCAATAGACC
	Rv	AGCTGCAGGCTTCTTACTAAACTCCTGTC
Peak_57185	Fw	ATCTCGAGGTGTAAGCTGCAGGGTCAAAG
	Rv	GCAGATCTTCAGCTTCCAGCACCTAAAAC
Peak_57188	Fw	CACTCGAGCTTAAGCTCAGTGCTGTTTCC
	Rv	TACTGCAGTTCCTAGCATCTGACACGGAG
Peak_83746	Fw	GACTCGAGCCTTTATGTCACAGGACCAAG
	Rv	TCCTGCAGGCTTCTGAGAGGTTAAGAGTC
Peak_92869	Fw	ACCTCGAGCACTTCGTTTGGTTAAAGCTG
	Rv	GCAGATCTGCCTTAACTGCCAAGCATAAC
Peak_102158	Fw	ACCTCGAGAGCTCCTTGCCTTTGATAAGC
	Rv	TTCTGCAGCCATCTTGCGAAGGTGACTGG
Peak_123246	Fw	GACTCGAGTCTCCTGCCTTGGATTTTCTG
	Rv	AGCTGCAGTCTGAAGCAGGAAAATGGACC

Whole-mount *in situ* hybridization

Whole-mount in situ hybridization was performed as previously described (Aires et al., 2019) using digoxigenin-labeled RNA antisense probes. For the genetically modified embryos and their wild type controls, at least 3 embryos were stained per probe and genotype, showing highly reproducible patterns. RNA probes have been previously described: *Msgn1* (Aires et al., 2019); *Sox2*, *Tbxt* and *Uncx4.1* (Aires et al., 2016); *Shh* and *Cdx2* (Jurberg et al., 2013); *Wnt5a* (Lickert et al., 2001) and *Pax3* (Goulding et al., 1991). Stained embryos

were included in 0.45% gelatin (Merck, 104078), 27% bovine serum albumin (Roche, 9048-46-8) and 18% sucrose (Sigma, S0389) in PBS and then polymerized with 1.75% glutaraldehyde (Biochem Chemopharma, 507130500) and sectioned at 35µm with a vibratome (Leica, VT1000S).

RT-qPCR

Total RNA was extracted from the caudal region of wild type and *Wnt5a*$^{\Delta CR1/\Delta CR1}$ embryos at E9.5 and E10.5 using Tri Reagent. 1 µg of RNA was used for reverse transcription into complementary DNA (cDNA) using NZY Reverse Transcriptase enzyme (NZYTech, MB124) and random hexamer mix (NZYTech, MB12901) following the manufacturer's protocol. Real-time qPCR was performed in a QuantStudio 7 Flex real-time PCR system (Thermo Fisher) using iQ SYBR Green Supermix (Bio-rad, 1708880) according to manufacturer's instructions. Primers used are listed in Table 2.5. Quantification was determined using the standard curve method, and expression levels normalized to *β-Actin*. Statistical significance was assessed using unpaired t-test.

Table 2.5 - Primers used in RT-qPCR.

Wnt5a	Fw	CCATGTCTTCCAAGTTCTTCC
	Rv	TACTTCTGACATCTGAACAGG
β-Actin	Fw	TCTGGTGGTACCACCATGTAC
	Rv	TACTTGCGCTCAGGAGGAGC

Acknowledgments

We thank the IGC animal house and Genomics Facilities for their expert services, advice and assistance. We also acknowledge Anastasiia Lozovska for the help with embryo culture and RT-qPCR data analysis, and all members of the Mallo laboratory for helpful discussions and comments throughout the course of the project.

This work was supported by the Fundação para a Ciência e a Tecnologia (LISBOA-01-0145-FEDER-030254) to M.M., by PhD fellowships from the Fundação para a Ciência e a Tecnologia (PD/BD/138240/2018 & COVID/BD152639/2022) to P.D., by the Fundação para a Ciência e a Tecnologia (PTDC-MEC-AND-30221-2017) and National Institutes of Health, National Library of Medicine Program (01LM011945-01) to R.B.C., and by the research infrastructure Congento (LISBOA-01-0145-FEDER-022170).

Chapter 3

Impact of retinoic acid signaling in the developing embryo

Part of chapters 2 and 3 are included in:

Duarte P., Correia R., Nóvoa A., Mallo M. Regulatory changes associated with the head to trunk developmental transition (2022) bioRxiv. doi.org/10.1101/2022.12.18.520961

Author contributions

Patrícia Duarte and Moisés Mallo designed experiments. Moisés Mallo generated *Raldh2* mutants. Patrícia Duarte performed experiments. Patrícia Duarte and André Dias performed single cell RNA sequencing. Ana Nóvoa performed pronuclear microinjection of DNA constructs. Patrícia Duarte and Moisés Mallo analyzed data. Patrícia Duarte wrote this chapter and Moisés Mallo edited it.

Introduction

Retinoic acid (RA) is a key signaling molecule during embryonic development. Heterodimers of nuclear receptors retinoic acid receptor (RAR) and retinoid X receptor (RXR), each with different isoforms, are bound to retinoic acid responsive elements (RAREs) (Benbrook et al., 1988; Brand et al., 1988; Dawson and Xia, 2012; Giguere et al., 1987; Zelent et al., 1989). Binding of RA leads to a conformational change in these heterodimers resulting in recruitment of either nuclear coactivators or corepressors, which in turn interact with histone acetylases and deacetylases, respectively, to regulate chromatin accessibility and transcription of the target genes (Chen and Evans, 1995; Dey et al., 2007; Hong et al., 2018; Kumar and Duester, 2014; Kurokawa et al., 1995).

Mice lacking RA signaling (*Raldh2* mutants) are developmentally arrested at the head to trunk transition, displaying problems in posterior neurogenesis and somitogenesis (Cunningham et al., 2015a; Niederreither et al., 1999). Indeed, the absence of RA-mediated repression of caudal *Fgf8* expression leads to expansion of the Fgf8 gradient along the PSM, culminating in smaller somites in *Raldh2*$^{-/-}$ (Cunningham et al., 2015a). Loss of RA activity is also accompanied by an imbalance in NMC cell differentiation, with decreased *Sox2*-expressing neural progenitors and increased *Tbx6*-expressing mesodermal progenitors (Cunningham et al., 2015a). In addition, RA is involved in the specification of neural progenitors to a motor neuron fate, where together with Shh, RA activates *Pax6* and *Olig2* expression in the spinal cord (Diez del Corral et al., 2003; Molotkova et al., 2005; Novitch et al., 2003). Interestingly, the axial truncation observed in *Raldh2* mutants can be rescued by an acute exogenous RA administration coincident with the timing of the head to trunk transition (Zhao et al., 2009). Embryos rescued in this manner are then able to complete trunk and tail development in the absence of further RA signaling activity, indicating that RA is only required for axial extension during the transition phase (Zhao et al., 2009).

The head to trunk transition is associated with substantial changes in chromatin accessibility of regulatory regions (as described in chapter 2). In this study we aimed to assess whether RA signaling is involved in the regulation of these extensive chromatin changes. For this, we compared chromatin accessibility profiles from the posterior epiblast region of wild type and *Raldh2*$^{-/-}$ mouse embryos at E8.5. We observed that the absence of RA activity had a limited impact on the changes observed in chromatin accessibility in the caudal epiblast at this developmental stage, and only a small fraction of the affected regions was associated with RAREs. In addition, we characterized the functional relevance of potential enhancers of *Nr2f2* and *Hoxa4*, identified in the accessibility analysis. Dissection of TF binding sites associated with each of these enhancers provided a valuable insight into their upstream regulation, revealing a complex interplay between elements for proper regulation of target gene expression. We also generated single cell transcriptome profiles of wild type and *Raldh2*$^{-/-}$ whole embryos at E8.25. We found a pronounced impact of the absence of RA activity in the caudal epiblast. Inference of directional trajectories (Bergen et al., 2020; La Manno et al., 2018) highlighted perturbations in the neural differentiation of progenitor cells from the caudal epiblast, consistent with the observed decrease of the spinal cord cell population in *Raldh2*$^{-/-}$ mutants.

Results & Discussion

Impact of RA signaling on chromatin accessibility

Genetic analyses revealed a fundamental role of RA signaling for proper transition from head to trunk development (Niederreither et al., 1999). We therefore tested the extent to which this is associated with changes in chromatin accessibility. Principal component analysis segregated samples by genotype (Fig 3.1A). Comparison of the global accessibility profiles from the posterior epiblast region of E8.5 wild type and *Raldh2*$^{-/-}$ mutant embryos revealed that only 120 peaks were differentially accessible between both conditions,

including 54 regions with decreased and 66 regions with increased accessibility in the *Raldh2*$^{-/-}$ mutants (Fig 3.1B). Moreover, we observed that the differentially accessible peaks mapped mainly to intergenic regions (43%) (Fig 3.1C), indicating that they might represent regulatory elements.

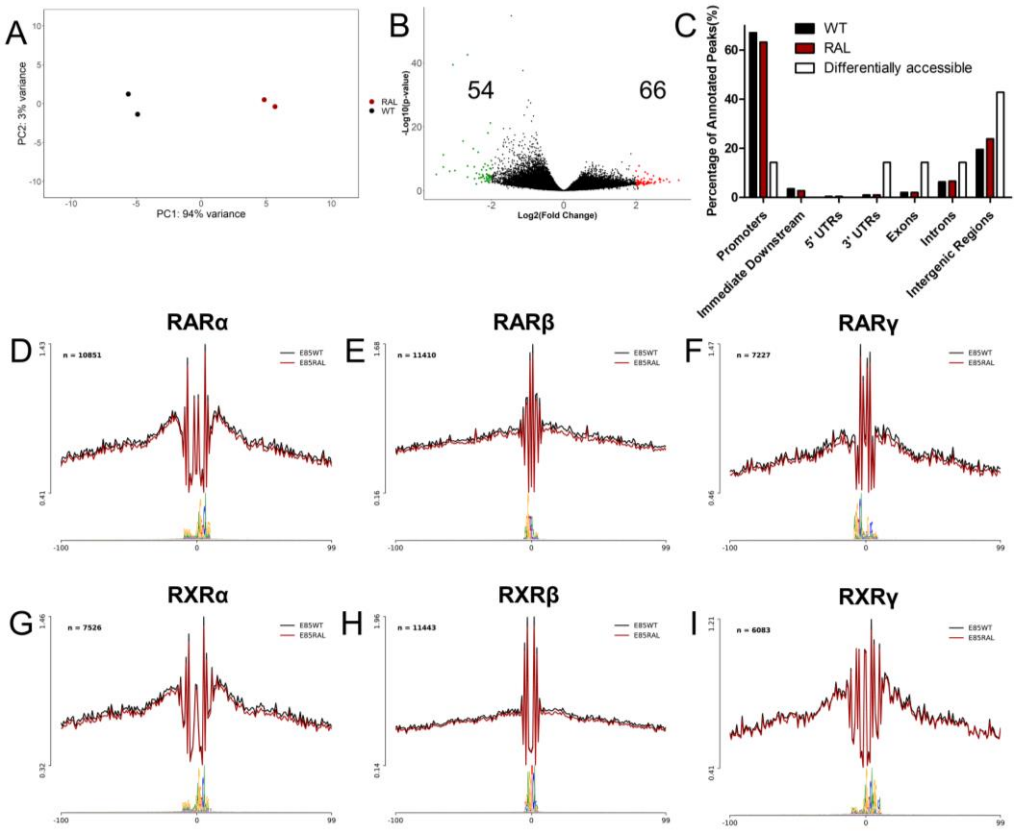


Figure 3.1 - Impact of RA signaling in genome accessibility of the posterior epiblast. (A) Principal component analysis of ATAC-seq data from WT (black) and *Raldh2*$^{-/-}$ (red). (B) Volcano plot of ATAC-seq peaks (*Raldh2*$^{-/-}$ vs WT) (|Log_2(Fold Change)|>2 & p-value < 0.05). Significantly open regions in *Raldh2*$^{-/-}$ are red, closed regions in *Raldh2*$^{-/-}$ are green and non-significant in black. (C) Genomic distribution of ATAC-seq peaks identified in WT (black), *Raldh2*$^{-/-}$ (red) and distribution of only the differentially accessible peaks (white). (D-I) Average ATAC-seq profiles of RARα (D), RARβ (E), RARγ (F), RXRα (G), RXRβ (H) and RXRγ (I) binding sites. Black profiles correspond to wild type, red profiles to *Raldh2*$^{-/-}$, n indicates the number of binding sites used to calculate the average profiles.

TF footprinting analyses showed no significant change in RAR and RXR binding activity in *Raldh2*$^{-/-}$ mutants (Fig 3.1D-I). This fits the notion that RA receptors are normally bound to retinoic acid response elements (RAREs) but kept inactive until bound by RA, (Kurokawa et al., 1995). Interestingly, only 12 of the regions that became differentially accessible contain RA receptor binding sites. Together, these observations suggest that RA activity in the posterior epiblast at this developmental stage does not involve major changes in the genomic regions bound by RA receptors and that most of the differences in chromatin accessibility observed in the *Raldh2* mutants are not mediated by direct RA activity, most likely representing instead downstream effects of factors under direct RA regulation. The potential involvement of genes regulated by the 12 elements containing binding sequences for RA receptors in this or other RA-dependent regulatory processes will require direct experimental analyses.

To further explore the differentially accessible chromatin regions identified in *Raldh2*$^{-/-}$ mutants, we performed additional functionality assays to evaluate their regulatory potential. For this we cross-referenced the identified 120 differentially accessible peaks with basewise conservation data (Pollard et al., 2010) and found that only 12 regions are phylogenetically conserved in mammals and therefore could have an evolutionarily conserved regulatory function. As such, we focused on these 12 elements and assessed their capacity to induce activity patterns. Of the 12 regions tested, five elements induced expression of the reporter gene in transgenic embryos (Table 3.1). Peak_14942 activated reporter expression restricted to the upper cleft of the first branchial arch (Fig 3.2A). Peak_123811 promoted expression in the PSM and caudal lateral epiblast (Fig 3.2B), while Peak_143897 showed activity in the PSM and most recently formed somites (Fig 3.2C). Peaks 135397 and 125080 generated prominent expression patterns and were the focus of further analyses, described in the sections below.

Table 3.1 - List of genomic regions tested in *β-galactosidase* reporter assays.

ID	Chr	Start coordinates	End coordinates	Regulation in RAL	Reporter expression
Peak_10773	chr1	188 884 705	188 885 217	Open	No
Peak_11325	chr1	194 058 650	194 059 176	Open	No
Peak_14942	chr10	68 248 668	68 248 949	Open	Yes
Peak_35261	chr12	97 402 840	97 403 128	Open	No
Peak_59634	chr16	25 843 726	25 844 047	Closed	No
Peak_60984	chr16	43 347 899	43 348 139	Open	No
Peak_67886	chr17	63 342 975	63 343 437	Open	No
Peak_123811	chr6	32 363 534	32 364 151	Open	Yes
Peak_125080	chr6	51 975 757	51 976 081	Closed	Yes
Peak_135397	chr7	70 847 893	70 848 657	Closed	Yes
Peak_143897	chr8	48 793 076	48 793 535	Open	Yes
Peak_152170	chr9	43 035 165	43 035 550	Open	No

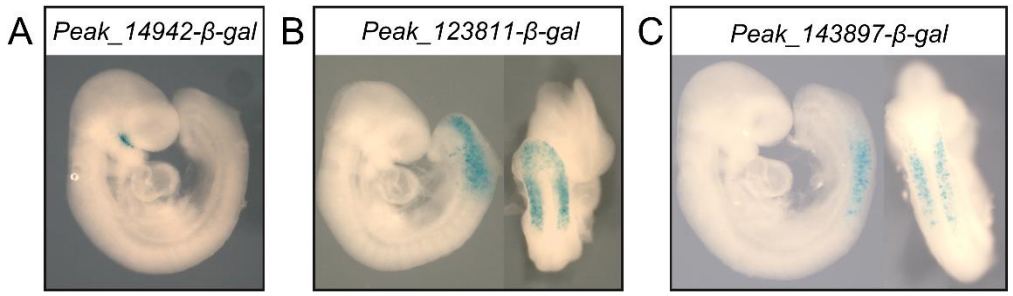


Figure 3.2 – Functionally active genomic regions. β-gal reporter expression in *Peak_14942-β-gal* (A), *Peak_123811-β-gal* (B) and *Peak_143897-β-gal* (C) transgenic embryos.

Evaluating potential *Nr2f2* enhancers for RA-mediated *Nr2f2* activation

Peak_135397 (we will refer to it as CR3) gained accessibility at E8.5 in a RA-dependent fashion (Fig 3.3A) and is located within the same TAD as *Nr2f2* (Fig 3.3B), a gene that is under the control of RA signaling (Berenguer et al., 2020). We selected CR3 for further analysis because when tested in transgenic reporter assays it reproduced to a large extent the *Nr2f2* expression pattern (Fig 3.3C-D), thus suggesting that it might be involved in the RA-dependent activation of *Nr2f2* expression. CR3 contains two distinct elements (CR3a and

CR3b) (Fig 3.3A). Both elements also gave well-defined activity profiles when tested individually in transgenic reporter assays. *CR3a-β-gal* embryos displayed staining in the somites starting from the forelimb level, in the second branchial arch neural crest (Fig 3.3E) and rhombomere 5 (Fig 3.3E''). CR3b gave a much broader range of expression in the neural tube, including the whole hindbrain and the spinal cord, and in the neural crest migrating from the hindbrain into the branchial arches (Fig 3.3F-F''). It also activated expression in the most anterior somites, where CR3a activity was not observed (Fig3.3E and F). Together, these staining patterns indicate that CR3 activity in the somites, branchial arches, and hindbrain might result from the combined CR3a and CR3b activities. However, the strong *CR3b-β-gal* reporter staining in the spinal cord contrasts with the absence of staining in the same region of CR3 reporter transgenics (Fig3.3 D-D'' and F-F''), suggesting that CR3a could block CR3b activity in this region.

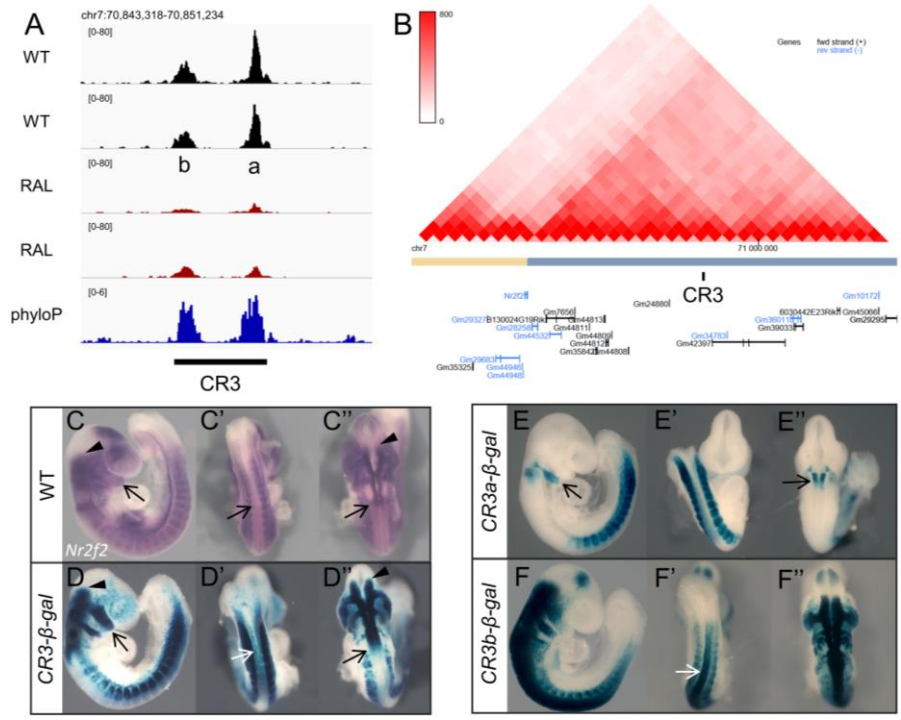


Figure 3.3 - Characterization of the CR3 region. (A) ATAC-seq tracks showing accessibility profiles in the CR3 region. CR3 includes two peaks, *a* and *b*. Phylogenetic conservation data

(phyloP)(Pollard et al., 2010) is shown in dark blue. (B) Hi-C data from 3D genome browser (Wang et al., 2018) highlighting CR3 location in the same TAD as *Nr2f2*. (C-D'') Comparison of *Nr2f2* expression pattern in wild type embryos by *in situ* hybridization (C-C'') with ***β-gal*** reporter expression in *CR3-β-gal* transgenic embryos (D-D''). Expression is present in the hindbrain (arrows in C'' and D'') but absent from the spinal cord (arrows in C' and D'). (E-E'') ***β-gal*** reporter expression in *CR3a-β-gal* transgenic embryos. Reporter expression is restricted to the second branchial arch (arrow in E), rhombomere 5 (arrow in E'') and somites from the forelimb level. (F-F'') ***β-gal*** reporter expression in *CR3b-β-gal* transgenic embryos is extended to the anterior somites and neural tube (arrow in F').

To further analyze the mechanisms regulating CR3 enhancer activity and the interactions between CR3a and CR3b, we searched for the presence of TF binding sites within these elements with HINT-ATAC. We identified two TF footprints in CR3a, matching Msgn1 and Hox binding sites (Fig 3.4A). Given the important role of these TFs in embryonic development, we assessed their contribution to CR3a enhancer activity by generating transgenic reporters for the CR3a element lacking each of these features. Transgenic embryos generated with CR3a lacking the Msgn1 binding site (*CR3a*$^{\Delta Msgn1}$) lost almost completely reporter gene expression in the somites (Fig 3.4B-B''), consistent with the known role of *Msgn1* as a regulator of paraxial mesoderm (Chalamalasetty et al., 2014; Yoon and Wold, 2000). Conversely, transgenic embryos of CR3a reporters lacking the Hox binding site (*CR3a*$^{\Delta HOX}$) did not affect somite expression, displaying instead extended reporter activity in the neural tube, including rhombomeres 3, 4 and 6 and the anterior spinal cord (Fig 3.4C-C''). This suggests a repressor rather than an activator role for Hox proteins in this enhancer, most particularly in the neural tube. We therefore tested whether the Hox binding site could also be involved in keeping CR3 inactive in the spinal cord by silencing CR3b activity in this embryonic region. Deletion of the Hox binding site from the CR3 reporter construct (*CR3*$^{\Delta Hox}$) resulted in a substantial activation of reporter activity in the neural tube (Fig 3.4D-D''), although not as extensive as the pattern obtained with CR3b (Fig

3.3F-F''), indicating that it could indeed be part of the interaction mechanism between CR3a and CR3b.

We also identified binding sites for Smad1 and Sp5 in the CR3b element (Fig 3.4E). Deletion of both sites resulted in the loss of reporter expression in most of the embryo, with some residual expression being detected in the hindbrain, neural crest and anterior spinal cord up until the trunk level (Fig 3.4F). Hence, the CR3a and CR3b elements are regulated by distinct sets of TFs, further allowing these regions to drive robust gene expression patterns despite possible fluctuations in upstream TF levels (Waymack et al., 2020).

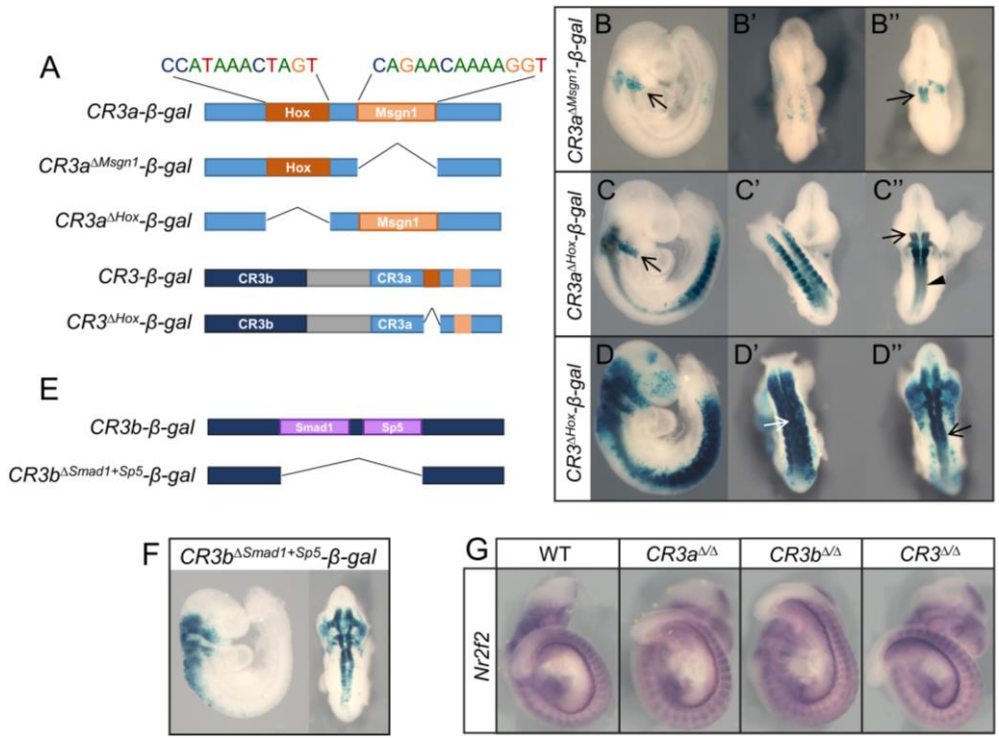


Figure 3.4 – The role of TFs in the regulation of CR3. (A) schematic representation of generated transgenic reporters for CR3a and CR3 regions lacking the specified TF binding sites. (B-B'') β-gal reporter expression in *CR3a$^{\Delta Msgn1}$-β-gal* is limited to the second branchial arch (arrow in B) and rhombomere 5 (arrow in B''). (C-C'') *β-gal* reporter expression in *CR3a$^{\Delta Hox}$-β-gal* transgenic embryos expands up to rhombomere 3 (arrow in C'') and along the neural tube (arrowhead in C''). (D-D'') *β-gal* reporter expression in *CR3$^{\Delta Hox}$-β-gal* transgenic embryos is extended into the neural tube (arrows). (E) Schematic representation of generated transgenic reporters for CR3b lacking the specified TF binding sites. (F) *β-gal* reporter expression in *CR3b$^{\Delta Smad1+Sp5}$-β-gal*

transgenics. (G) Whole mount *in situ* hybridization of wild type, *CR3a*$^{\Delta/\Delta}$, *CR3b*$^{\Delta/\Delta}$ and *CR3*$^{\Delta/\Delta}$ at E9.5 using a probe for *Nr2f2*.

Together, the reporter assays indicate the existence of regulatory interactions between the CR3a and CR3b elements to achieve a pattern of activity resembling *Nr2f2* expression. CR3 thus represents a case in which enhancer interactions, both positive and negative, play an important role in fine tuning gene expression, contributing to the production of sharp boundaries in the expression domains, similarly to what has been previously reported for other systems (Bothma et al., 2015; El-Sherif and Levine, 2016; Perry et al., 2011). Our results also suggest that the RA-dependent opening of CR3a and CR3b might expose these elements to become activated by factors involved in the development of trunk and hindbrain structures.

To directly assess CR3 function and its potential relevance for *Nr2f2* expression, we generated mouse strains containing deletions of CR3a, CR3b and CR3. Homozygous mutant animals for each of these strains developed to term and the adults had no obvious phenotypic alterations, already indicating that these mutants kept *Nr2f2* expression, at least to a level allowing normal development. We confirmed this by whole-mount *in situ* hybridization showing that the *Nr2f2* expression pattern in homozygous mutant embryos for any of the deleted CR3 regions were similar to that observed in wild type embryos (Fig 3.4G). This indicates that if CR3a and CR3b are indeed involved in *Nr2f2* expression as suggested by the reporter assays, other redundant enhancers might be present that ensure *Nr2f2* expression and prevent developmental arrest caused by the inactivation of this gene. A possible candidate for an enhancer able to maintain *Nr2f2* transcription in the absence of CR3 is the previously identified RARE in intron 1 of *Nr2f2* (Berenguer et al., 2020). Consistent with this hypothesis, we observed that the genomic region featuring this RARE is accessible in both our wild type and *Raldh2* mutant datasets. However, further studies will be required to validate the role of this RARE in the regulation of *Nr2f2* and whether it interacts functionally with CR3.

Assessing a potential enhancer of *HoxA* genes

Another candidate region that was further analyzed for its regulatory potential was Peak_125080 (we will refer to it as CR4). CR4 consistently activated reporter expression in the neural tube at E9.5, restricted to the trunk region (Fig 3.5A). CR4 was open and accessible in wild type embryos at E8.5, whereas no ATAC-seq peak was observed in *Raldh2*$^{-/-}$ mutants (Fig 3.5B). Hence, in the absence of RA signaling this region failed to become accessible and could thus represent an enhancer region under the control of RA signaling. Identifying the target gene of CR4 is a crucial step in understanding its regulatory function. CR4 maps to intron 4 of *Skap2*, which is located in the same TAD as the *HoxA* cluster (Fig 3.5C). In addition, intron 4 of *Skap2*, has been shown to engage in strong interactions with genomic regions from *Hoxa4* to *Hoxa7* (Gentile et al., 2019). This suggests that CR4 may regulate a *HoxA* gene, more specifically either *Hoxa4, Hoxa5, Hoxa6* or *Hoxa7*.

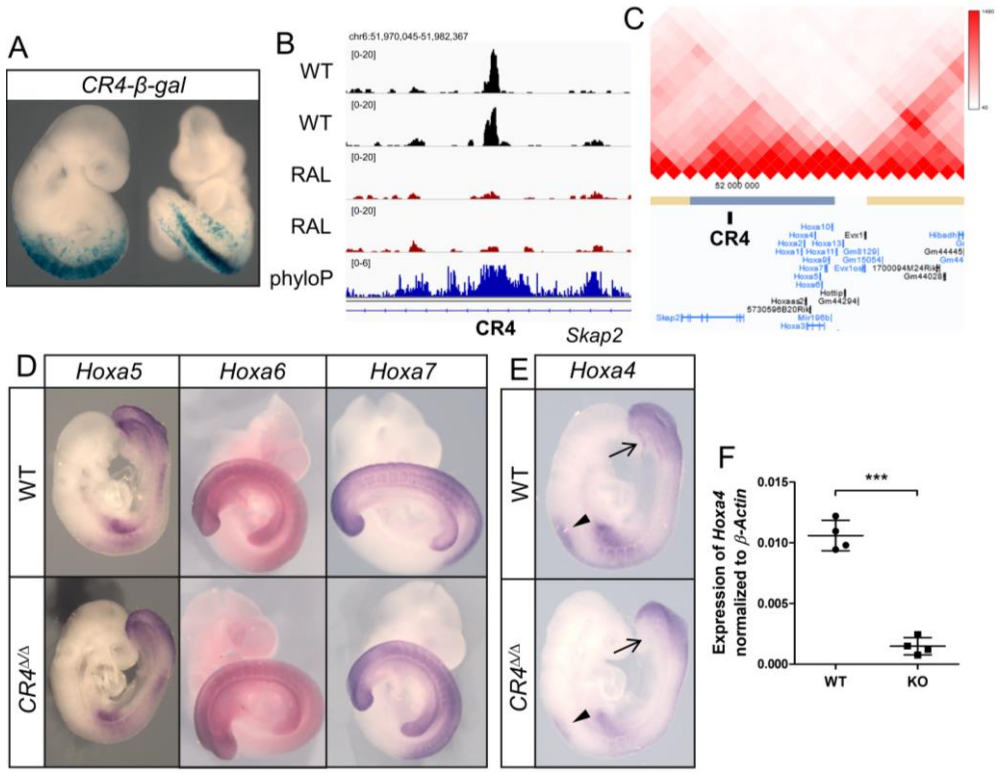


Figure 3.5 - Characterization of the CR4 region. (A) *β-gal* reporter expression in *CR4-β-gal* transgenic embryos. (B) ATAC-seq tracks showing accessibility profiles in the CR4 region. Phylogenetic conservation data (phyloP)(Pollard et al., 2010) is shown in dark blue. (C) Hi-C data from 3D genome browser (Wang et al., 2018) highlighting CR4 location in the same TAD as the HoxA cluster. (D) Whole mount *in situ* hybridization of wild type and $CR4^{\Delta/\Delta}$ embryos at E9.5 using a probe for *Hoxa5, Hoxa6 and Hoxa7*. (E) Whole mount *in situ* hybridization of wild type and $CR4^{\Delta/\Delta}$ embryos at E9.5 using a probe for *Hoxa4.* Arrows and arrowheads emphasize reduced expression of *Wnt5a* in the caudal region and spinal cord, respectively. (F) RT-qPCR analysis of *Hoxa4* gene expression in wild type (WT) and $CR4^{\Delta/\Delta}$ (KO) embryos at E9.5. *Hoxa4* expression is normalized to *β-Actin*. Error bars indicate the standard deviation. ***, p-value < 0.001.

To directly test this hypothesis, we generated CR4 deletion mutants ($CR4^{\Delta/\Delta}$). Whole-mount *in situ* hybridization revealed that the patterns of expression for *Hoxa5, Hoxa6* and *Hoxa7* were similar in wild type and CR4 deletion mutant embryos (Fig 3.5D). However, an overall reduction of *Hoxa4* expression levels was evident in $CR4^{\Delta/\Delta}$ at E9.5, confirmed by quantitative RT-PCR analysis (Fig 3.5E and F), suggesting that CR4 might act as a *Hoxa4* enhancer. Even though deletion of the CR4 element had an obvious quantitative effect on *Hoxa4* expression, transcripts for this gene were still observed in $CR4^{\Delta/\Delta}$ mutants, likely induced by redundant enhancers. Indeed, RA signaling has previously been described to activate *Hoxa4* expression through a RARE in the 5' flanking region of the *Hoxa4* promoter (Behringer et al., 1993; Packer et al., 1998). This, combined with *Hox4* redundant paralog function (Horan et al., 1995), might account for why CR4 homozygous mutants are viable and adult animals show no apparent phenotype.

Next, we searched for the presence of TF binding sites within CR4 with HINT-ATAC to explore the mechanisms regulating its activity. With this analysis we identified three distinct TF footprints, whose motifs matched to Rfx6, Pax3 and Vdr binding sites respectively. To evaluate their role in CR4 enhancer activity, we generated transgenic reporters for the CR4 element lacking these binding sites (Fig 3.6A). Transgenic embryos with CR4 lacking both the Rfx6 and Pax3 binding sites ($CR4^{\Delta Rfx6+Pax3}$) displayed a pattern of reporter

expression in the neural tube (Fig 3.6B), similar to that observed in the *CR4-β-gal* transgenics (Fig 3.5A). However, there was increased reporter expression in the lateral mesoderm (Fig 3.6B), suggesting that one or both TFs could act as inhibitors in this tissue. In contrast, transgenic embryos with CR4 lacking the Vdr binding site (*CR4*$^{\Delta Vdr}$) completely lost reporter expression (Fig 3.6C), revealing that Vdr is essential to induce CR4 enhancer activity. A chromatin immunoprecipitation (ChIP) assay with an anti-Vdr antibody, followed by qPCR confirmed that Vdr binds to the putative binding site identified within the CR4 element at E8.5 (Fig 3.6D). Interestingly, Vdr forms heterodimers with retinoid X receptor (RXR) to bind specific DNA sequences and regulate target genes (Haussler et al., 2013). Overall, this indicates that Vdr binds and activates the CR4 regulatory region, in a RA dependent manner, which in turn will induce the expression of *Hoxa4*.

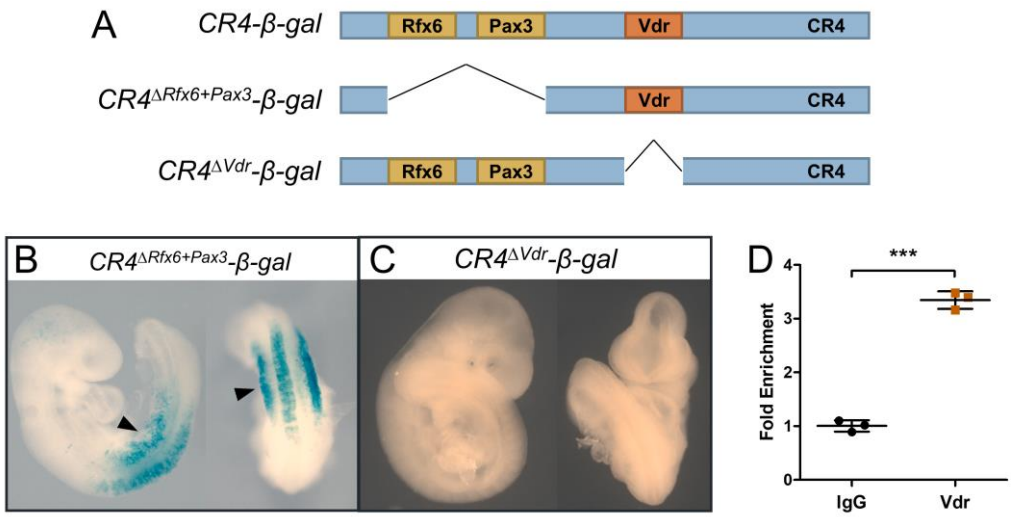


Figure 3.6 – Role of TFs in the regulation of CR4. (A) schematic representation of generated transgenic reporters for CR4 lacking the specified TF binding sites. (B) *β-gal* reporter expression in *CR4*$^{\Delta Rfx6+Pax3}$*-β-gal* transgenic embryos. Increased reporter expression is observed in the lateral mesoderm (arrowheads). (C) Absence of *β-gal* reporter expression in *CR4*$^{\Delta Vdr}$*-β-gal* transgenic embryos. (D) ChIP-qPCR of Vdr binding to CR4 in the posterior epiblast of wild type E8.5 embryos. Fold Enrichment was normalized to IgG (negative control), ***, p-value<0.001.

Transcriptome profile of the posterior epiblast of *Raldh2* mutants

To explore the impact of the absence of retinoic acid signaling we characterized the transcriptional profile of *Raldh2* mutants by RNA-seq. For this, we dissected the posterior epiblast region of mouse wild type (WT) and *Raldh2*$^{-/-}$ (RAL) embryos at E8.5, equivalent to the ATAC-seq experimental setup. Principal component analysis separated the samples by genotype (Fig 3.7A). Differential analysis revealed 372 genes that were significantly upregulated, and 948 genes that were downregulated in RAL, when compared with WT (Fig 3.7B). The observation that about three quarters of the differentially expressed genes were downregulated may allude to the activation of gene expression having a higher weight in the RA-dependent functions at this developmental stage. Intriguingly, manual inspection revealed that almost 52% of the total DEGs pertain to lncRNAs and unannotated genes (e.g. 4930517O19Rik, Gm10388), suggesting that several genes that might play important roles downstream of RA signaling still remain uncharacterized. Future studies geared towards generation of mice with deletions for some of these unannotated genes would be interesting to explore their function and possibly uncover new insights into the role of RA signaling at this developmental stage.

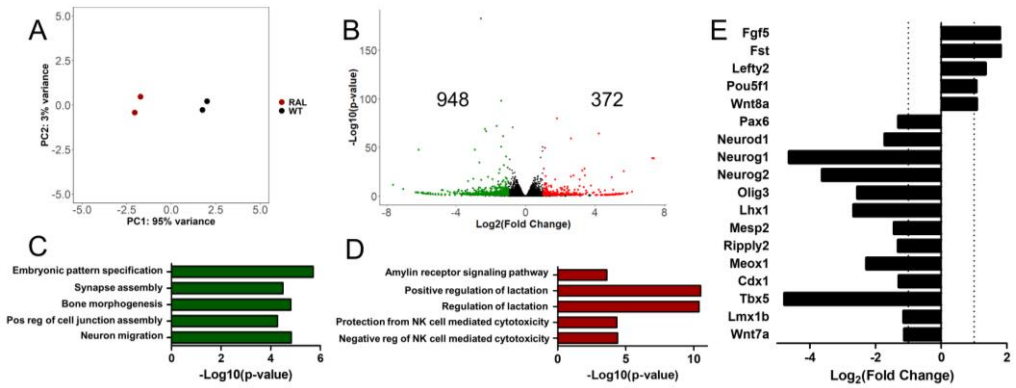


Figure 3.7 - Transcriptome profile of the posterior epiblast of *Raldh2*$^{-/-}$. (A) Principal Component Analysis of RNA-seq data from WT (black) and RAL (red) at E8.5. (B) Volcano plot of RNA-seq gene expression ($|Log_2(Fold\ Change)| \geq 1$ & p-value < 0.05). Significantly upregulated genes in RAL in red, downregulated genes in RAL in green and non-significant in black. (C) Top 5 enriched

GO biological process terms of downregulated genes in RAL. (D) Top 5 enriched GO biological process terms of upregulated genes in RAL. Abbreviations: Pos, positive; reg, regulation. (E) Expression levels of key developmental genes.

Even considering the limitations imposed by the large proportion of unannotated genes, we wanted to obtain an overview on the pathways affected by the absence of RA by exploring enrichment for GO terms among annotated genes within the DEGs (Fig 3.7C and D). For the downregulated genes, we observed enrichment of GO terms related to embryonic pattern specification and neural development (i.e. synapse assembly, neuron migration) (Fig 3.7C), consistent with known developmental requirements of RA signaling (Molotkova et al., 2005; Niederreither et al., 1999; Ribes et al., 2008; Zhao et al., 2009). These included genes like *Neurog2* and *Pax6*, involved in initiating and regulating neurogenesis, respectively (Lacomme et al., 2018; Ribes et al., 2008; Sansom et al., 2009), both of which were downregulated in our mutant datasets (Fig 3.7E). Also downregulated in RAL tissue were genes associated with specification of somitic mesoderm, including *Mesp2* and *Ripply2*, both required for somite segmentation by establishing the anterior boundary of *Tbx6*, (Takahashi et al., 2010; Zhao et al., 2015). *Meox1* and *Cdx1* were also downregulated in the RAL mutants, consistent with previous reports indicating their activation by RA signaling (Houle et al., 2003, 2000; Kennedy et al., 2009). *Tbx5*, *Lmx1b* and *Wnt7a* were also among the downregulated genes in the RAL embryos, which might reflect the requirement of RA signaling for the induction of forelimb bud formation (Adamska et al., 2005; Agarwal et al., 2003; Krawchuk and Kania, 2008; Zhao et al., 2009).

Upregulated genes in RAL embryos included early epiblast markers *Fgf5* and *Wnt8a*, and the pluripotent gene *Pou5f1*. Both *Pou5f1* and *Wnt8a* have been shown to be repressed by RA signaling (Cunningham et al., 2015b; Gupta et al., 2008; Pikarsky et al., 1994), and *Wnt8a* has been reported to collaborate with *Wnt3a* to maintain *Fgf8* expression and suppress *Sox2* in the caudal epiblast (Cunningham et al., 2015b). We also observed upregulation of *Fst* in

the caudal epiblast at a developmental stage when *Fst* expression is normally limited to the hindbrain and somites (Cunningham et al., 2016; Stafford et al., 2014). Altogether, these observations highlight how both neural and mesodermal specification are compromised in the absence of RA signaling.

GO analysis of the upregulated genes uncovered an enrichment for terms associated with very diverse biological processes that have not been previously linked directly to RA activity in the embryo (Fig 3.7D). One such process was the Amylin receptor signaling pathway, known to be involved in controlling energy and glucose homeostasis (Turek et al., 2010; Boccia et al., 2020). From the 7 genes in this GO category, three were found upregulated in our datasets (*Adm, Calcr* and *Ramp1),* which accounts for the high enrichment observed in the GO analysis. From these, *Calcr* had already been shown to be repressed by RA signaling (Lanigan et al., 1993; Bi et al., 2018). Another enriched GO term was regulation of lactation, with several prolactin genes found upregulated in RAL. Prolactin is one of the most abundant soluble factors present in the amniotic fluid, originating from both maternal and fetal production and is involved in regulation of reproduction, development, behavior, metabolism, osmotic balance, and immunity (Bole-Feysot et al., 1998; WINTERS et al., 1975). Interestingly, RA has been shown to inhibit expression of the prolactin receptor (Widschwendter et al., 1999), already supporting a direct role of RA in inhibiting prolactin signaling. Finally, we also observed enrichment in the GO term associated with regulation of natural killer cell mediated cytotoxicity. Indeed, in the absence of RA we observed an upregulation of genes that suppress natural killer cell cytotoxicity, including *Lgals9* and the *Serpinb9* gene family (El Haddad et al., 2011; Golden-Mason et al., 2013), consistent with a role for RA in the inhibition of natural killer cell cytotoxicity by decreasing the production of IFN-γ (Abb et al., 1982; Chang and Hou, 2015; Li et al., 2007). Overall, GO analyses of DEGs highlight the role of RA signaling in both promoting and inhibiting a wide range of systemic functions.

Single-cell transcriptome profile of *Raldh2* mutants

To have a better understanding of the early effects that absence of RA signaling has on the developing embryo at a cellular level we generated single cell RNA-seq data from both wild type and *Raldh2*$^{-/-}$ whole embryos at E8.25. Unsupervised clustering was applied independently to segregate each dataset into individual clusters. We then annotated the clusters by comparing the most expressed genes in each cluster (Fig 3.8 and 3.9) with known gene expression patterns in embryos, as well as with published scRNA-seq datasets (Pijuan-Sala et al., 2019). With this approach we were able to identify clusters representing the main tissues found in embryos at this developmental stage (Fig 3.10A). Even though the projection yielded similar clusters in RAL, we observed some differences (Fig 3.10B). In particular, the absence of an individual spinal cord cluster in RAL, although a small group of cells included in the hindbrain cluster present markers pertaining to the spinal cord. In addition, we identified two individual clusters (6 and 14) in the RAL dataset, both containing gene markers characteristic of the caudal epiblast (Fig 3.9), an embryonic region represented by a single cluster in WT embryos (Fig 3.10A).

To understand if this was a real separation or just an artifact resultant of the dimensionality reduction step, we analyzed the differential expression of cluster 14 *versus* cluster 6 (Fig 3.10C). This analysis revealed 30 genes differentially expressed between these two clusters, several of which are regulated by RA signaling. In cluster 6 we observed higher expression of *Xist* which is indirectly regulated by RA-dependent repression of *Pou5f1* (Ahn and Lee, 2010), as well as *Fst*, *Fos* and *Jun*, all repressed by RA (Cunningham et al., 2016; Lee et al., 1998; Ozeki and Tsukamoto, 1999). In cluster 14 we identified *Cdx1* which is activated by RA (Houle et al., 2003, 2000, p. 200), *Crabp2*, that encodes for a protein that transports RA from the cytosol to RARs present in the nucleus (Dong et al., 1999), and *Mdk* a RA responsive gene involved in neurite growth (Kaname et al., 1993). Altogether, segregation of clusters 6 and 14 may indicate a differentially regulated impact of the absence of RA signaling in the caudal epiblast.

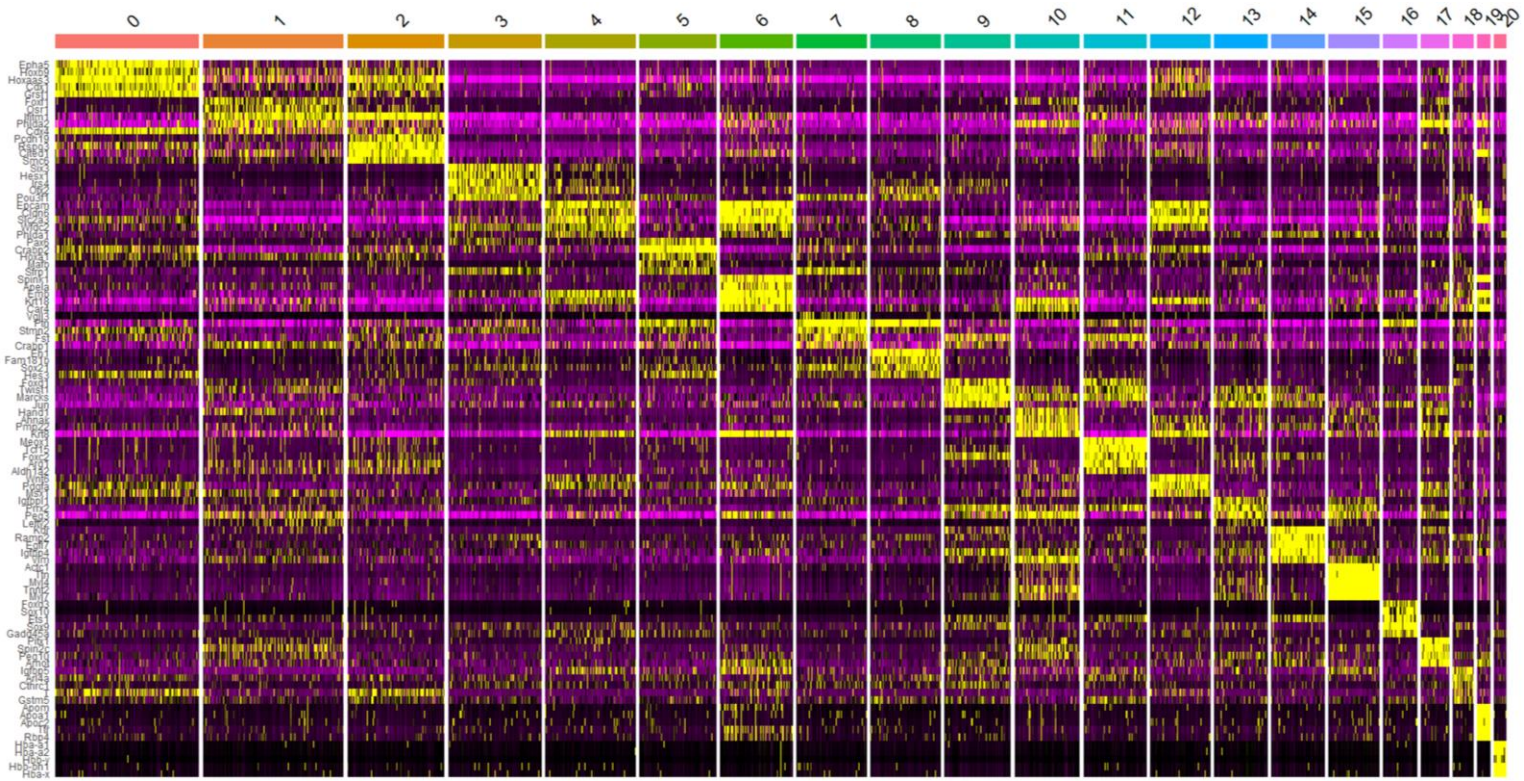


Figure 3.8 – Top 5 marker genes in wild type scRNA-seq dataset. Heatmap of the five most differentially expressed genes of each cluster in WT dataset. Color scale represents scaled gene expression (z-score values). Yellow represents higher expression levels, purple lower expression, and in black are the non-differentially expressed genes.

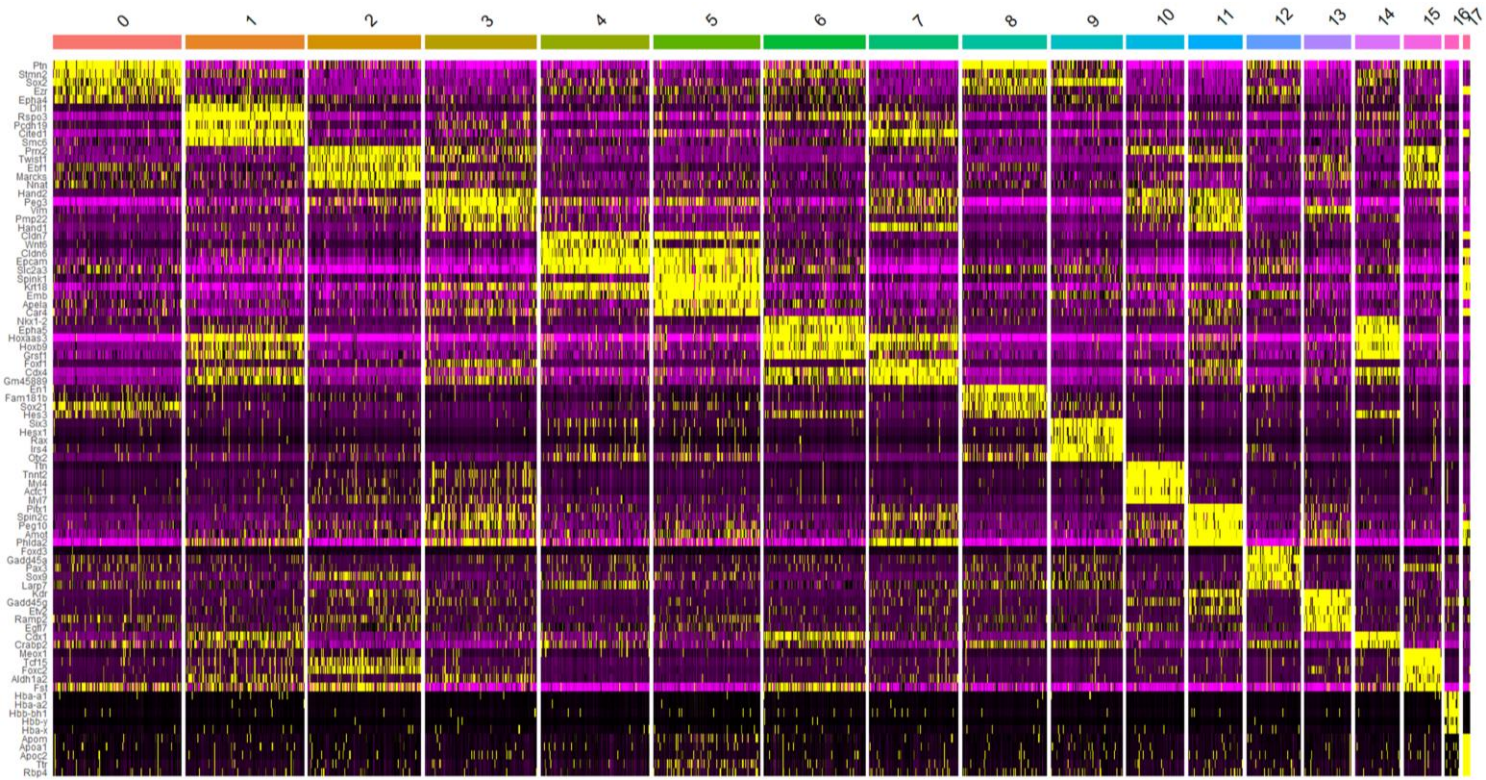


Figure 3.9 - Top 5 marker genes in *Raldh2* mutant scRNA-seq dataset. Heatmap of the five most differentially expressed genes of each cluster in RAL dataset. Color scale represents scaled gene expression (z-score values). Yellow represents higher expression levels, purple lower expression, and in black are the non-differentially expressed genes.

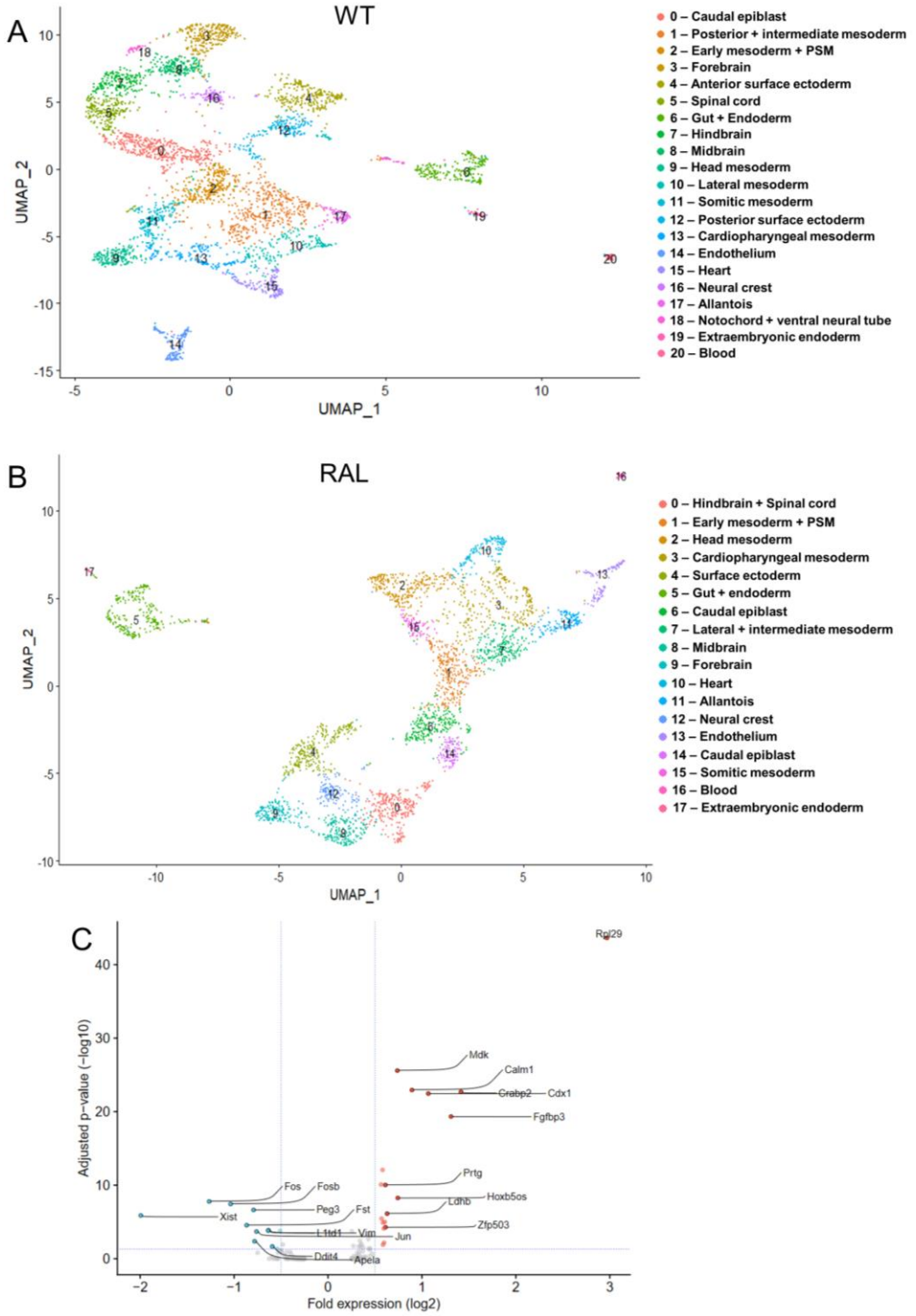


Figure 3.10 – Characterization of wild type and *Raldh2*$^{-/-}$ cell populations. UMAP clustering of wild type (A) and *Raldh2*$^{-/-}$ (B) whole embryo datasets. (C) Differential expression analysis

between the cluster 14 versus 6 for the *Raldh2*$^{-/-}$ dataset. In red are the genes more expressed in cluster 14 cells, in blue the genes more expressed in cluster 6 cells. The top 10 up- and downregulated genes are highlighted. Adjusted p-value <0.05 and $|Log_2(\text{Fold change})|>0.5$.

Next, we explored the lineage relationships in the WT and RAL single-cell datasets using scVelo, which relies on computation of splicing kinetics to infer directional trajectories (Bergen et al., 2020). The inferred trajectories along the WT clusters highlight already known lineages, including the emergence of the somitic mesoderm from the PSM, and the contribution of the lateral plate and cardiopharyngeal mesoderm to the heart (Fig 3.11A). Importantly, the trajectories in the WT UMAP reflected the duality of the caudal epiblast (cluster 0) in providing cells both for neural (cluster 5) and mesodermal tissues (cluster 2). However, in the RAL dataset the directional flows originating from the caudal epiblast clusters (6 and 14) were severely disrupted, lacking trajectories contributing to neural specification (Fig 3.11B). Indeed, from cluster 14 we only observed directional flows to cluster 6, whereas the flow from cluster 6 was only directed to mesodermal fates (cluster 1). These observations highlight the existence of perturbations in the neural differentiation of progenitor cells from the caudal epiblast in *Raldh2*$^{-/-}$ mutants already suggested in the transcriptomic assay.

To understand what could be causing this discrepancy in the inferred lineages, we identified the genes with the highest differential RNA velocity. Since these are the genes contributing the most to the inferred trajectories and could therefore be potential drivers of cell specification (Bergen et al., 2020). Comparison of the generated list of the top 100 dynamic genes for clusters 0 (WT), 6 and 14 (RAL) showed more genes in common between both RAL clusters than with WT (Fig 3.11C). The unique genes in WT could represent key driver genes that are compromised in the gene regulatory networks of *Raldh2* mutants therefore, we focused on the top 15 genes with highest differential velocity in each cluster. With this approach we identified 5 driver genes unique to the WT cluster (Table 3.2). Among these, *Fat3* might be

particularly relevant, as its conditional deletion in the neural tube has recently been shown to lead to depletion of neural progenitors, a consequence of the requirement of *Fat3* for the maintenance of progenitor proliferation (Seo et al., 2022). Together, this suggests that *Fat3* might also have a key regulatory role in maintaining progenitor proliferation in the caudal epiblast. Interestingly, we found that *Fat3* is downregulated in the posterior epiblast of *Raldh2*$^{-/-}$ at E8.5 (Log_2(Fold Change) = -1,54), and a similar downregulation has also been reported in the trunk region of *Raldh2*$^{-/-}$ at the same developmental stage (Berenguer 2020). Other unique driver genes identified in this analysis include *Rspo3*, which activates canonical Wnt signaling (Lebensohn and Rohatgi, 2018); *Gabrb2*, that encodes a subunit of $GABA_A$ synaptic receptor, with homozygous mutant mice presenting a schizophrenia-like phenotype (Yeung et al., 2018); *Gad1*, essential for the synthesis of GABA, inactivation of this gene produces lethal phenotypes at birth (Asada et al., 1997); and *Scg5* involved in regulation of pituitary hormone secretion (Mbikay et al., 2001).

Table 3.2 - Top 15 driver genes in WT cluster 0 and RAL clusters 6 and 14. Genes listed are ranked by differential RNA velocity values. Unique genes for each cluster are highlighted in bold.

WT Cluster 0	RAL Cluster 6	RAL Cluster 14
Hoxb6	*Lrrtm4*	*Hoxb6*
Arid3b	*Hoxb6*	*Lrrtm4*
Gabrb2	*B230323A14Rik*	*B230323A14Rik*
Hoxa9	*Dgkb*	*Tmem132c*
Hoxb5os	*Tmem132c*	*Spock3*
Spock3	*Hoxb3os*	*Hoxb5os*
Hoxd3	*Tbxt*	*Plpp4*
Rspo3	*Alcam*	*Tfcp2l1*
Kif26b	*Lypd6b*	*Sema3e*
Tmem132c	*Mtmr7*	*Alcam*
Gad1	*Sema3e*	*Mcc*
Fat3	*Ttc28*	***Oxr1***
Scg5	*Mcc*	*Hoxa9*
Slco5a1	*Hoxb5os*	*Ccdc171*
Sema3e	*Hmga2*	*Wnt3a*

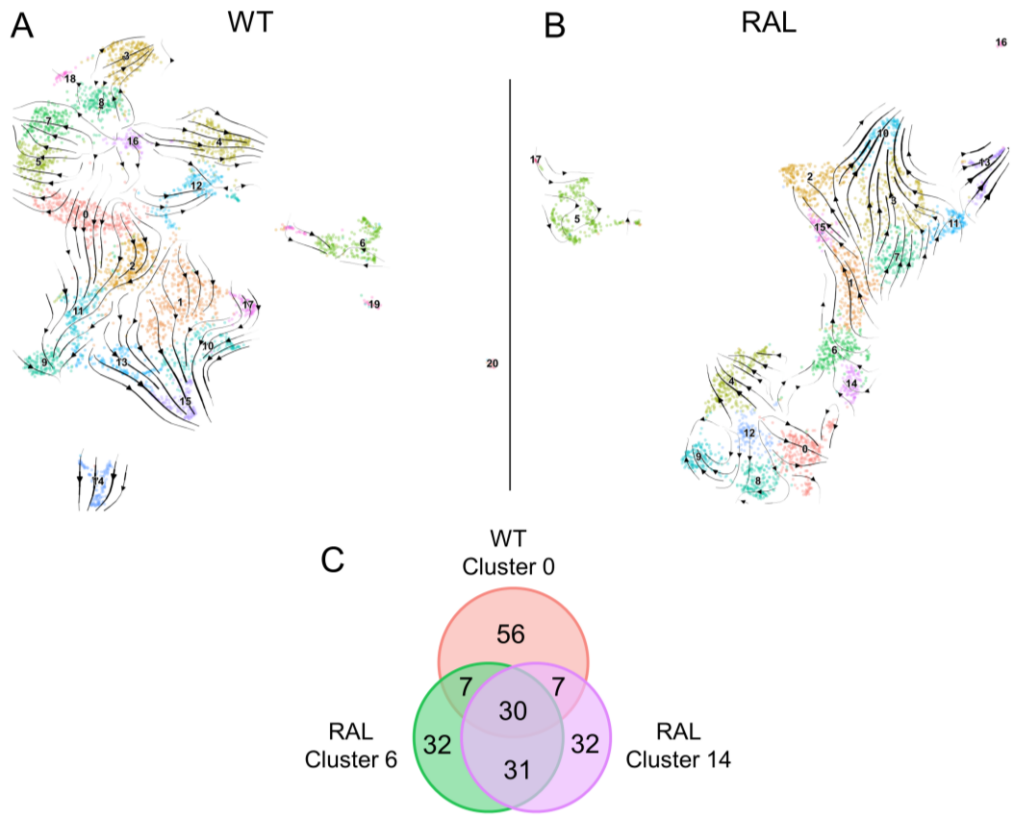


Figure 3.11 – Lineage relationship reveals differential trajectories in caudal epiblast cell populations. (A-B) Inferred directional trajectories in wild type (A) and *Raldh2*$^{-/-}$ (B) cell populations. Cluster identification according to Figure 3.10. (C) Venn diagram illustrating the overlap between the top 100 genes with the highest differential RNA velocity in WT cluster 0 and RAL clusters 6 and 14.

Direct comparison of separate single-cell datasets requires data integration, by identifying and aligning shared cell populations (Stuart et al., 2019). Following this approach, we used the WT as the baseline dataset to which RAL cells were mapped. Integrative analysis revealed the existence of cells in the RAL dataset aligning to every single cluster present in the WT UMAP (Fig 3.12A). However, the spinal cord cluster was found under-represented in *Raldh2*$^{-/-}$ mutants, comprising only 1.3% of total cells as opposed to the 4.9% found in WT (Fig 3.12B), which may explain why this cell population did not show as an individual cluster in the initial RAL UMAP. The marked

decrease of the spinal cord cell population is also consistent with the disruptions we observed in lineage trajectories toward neural fates in RAL.

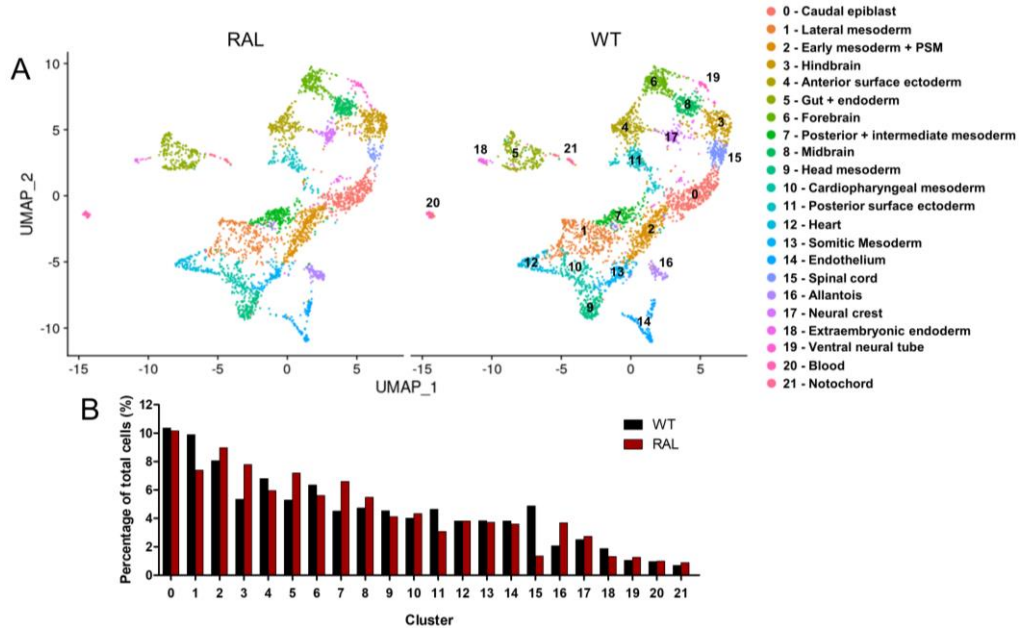


Figure 3.12 – Integration of wild type and *Raldh2*$^{-/-}$ datasets reveal an under-representation of spinal cord cell population. (A) UMAP clustering of the integrated datasets. (B) Percentage of total cells per cluster in wild type (black) and *Raldh2*$^{-/-}$ (red).

Comparative scRNA-seq analysis revealed significant expression changes in diverse systemic cell functions. We observed an overall upregulation of several *Atp, Cox* and *Nduf* genes (Fig 3.13A), which are an integral part of the electron transport chain and ATP synthase. The significance of this observation is not clear, as there have been conflicting reports about the impact that addition of *all trans* retinoic acid has on oxidative phosphorylation, in a cell culture context (Papa et al., 2017; Tourniaire et al., 2015). More recently, analysis of *Rdh10*$^{+/-}$ mice indicated a sexually dimorphic effect of RA on muscle metabolism, specifically in terms of complex IV activity and ATP production (Zhao et al., 2021). Moreover, we observed quantitative variations in the expression of various ribosomal genes. While several *Rpl* and *Rps* genes were found upregulated, *Rpl29* was found highly downregulated in RAL (Fig 3.13A).

Rpl29 is part of the 60S large ribosomal subunit and has been shown to be involved in regulating the rate of protein synthesis with significant functional impact, as homozygous mutants presented global growth defects and a postnatal survival rate of 50% (Kirn-Safran et al., 2007). *Fos* and *Jun,* key components of the JNK pathway, were also upregulated in RAL (Fig 3.13B and C), fitting with previously reported RA repressive functions (Lee et al., 1998; Ozeki and Tsukamoto, 1999). Interestingly, these were two of the genes that led to the split of the caudal epiblast into two individual clusters (Fig 3.10C).

Although not many developmental genes were found differentially expressed at this stage, *Raldh2*$^{-/-}$ mutants showed increased expression levels of *Wnt5a* and *Wnt8a* (Fig 3.14A and B)*,* both required for axial extension and maintenance of progenitor cells (Cunningham et al., 2015b; Yamaguchi et al., 1999). We could also observe an overall upregulation of *Fgf8, Fst, Pou5f1* and *Tbxt* (Fig 3.14C-F), consistent with their RA dependent regulation (Cunningham et al., 2016; Gupta et al., 2008; Kumar and Duester, 2014; Martin and Kimelman, 2010). Since scRNA-seq was performed on embryos at an earlier developmental stage, it might be too early for the appearance of some of the differences observed in the epiblast at E8.5 by RNA-seq (Fig 3.7). However, it provides an insight into which are the first developmental genes affected by the absence of RA signaling. Interestingly, these are mostly genes that require RA-dependent repression during normal development, leading to extended domains of expression along the embryo.

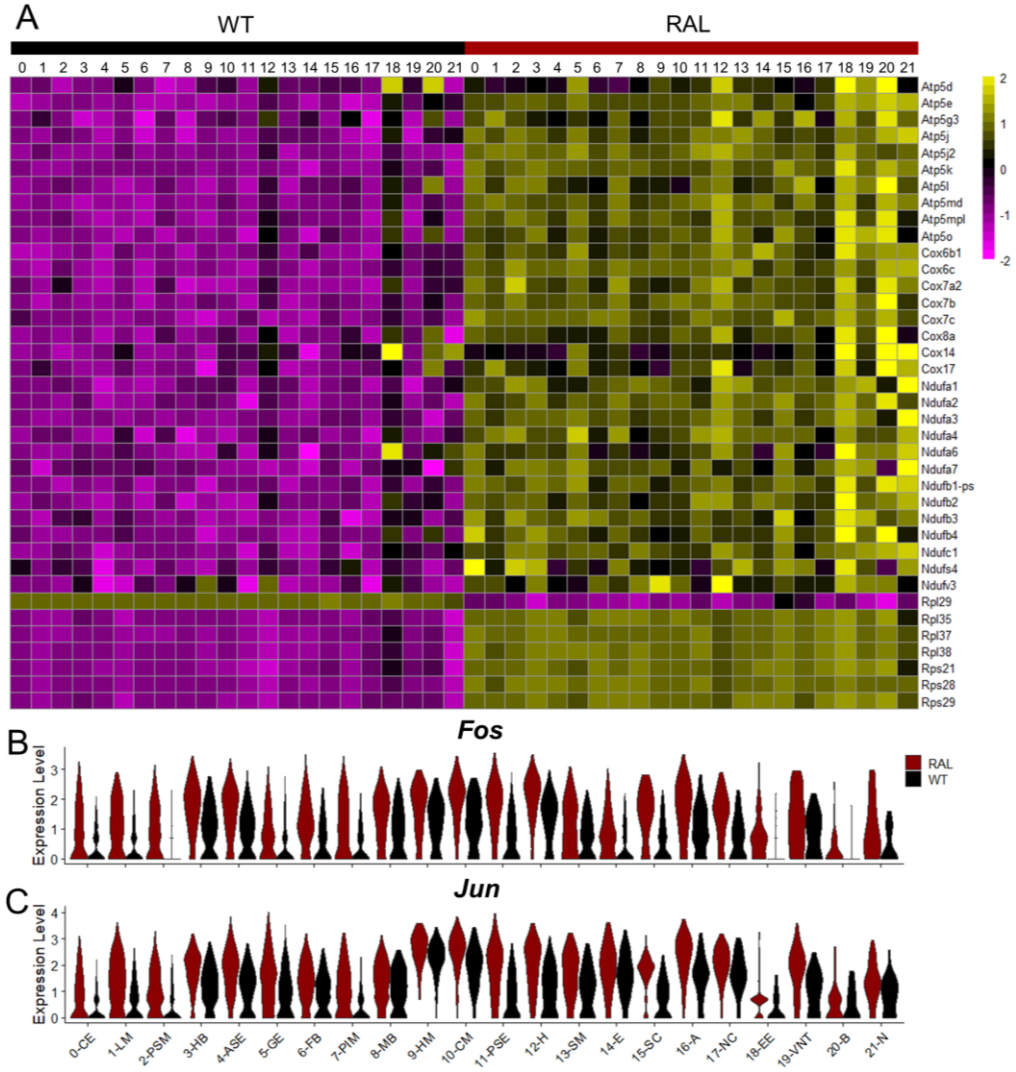


Figure 3.13 – Absence of RA signaling leads to global changes in gene expression. (A) Heatmap comparing average expression of select genes in each cluster of the integrated wild type and *Raldh2*$^{-/-}$ datasets (as identified in Figure 3.12A). Color scale represents scaled gene expression (z-score values). (B-C) Expression levels of *Fos* (B) and *Jun* (C) in each cluster of the integrated wild type (black) and *Raldh2*$^{-/-}$ (red) datasets (as identified in Figure 3.12A). Abbreviations: CE, Caudal epiblast; LM, Lateral mesoderm; PSM, Presomitic mesoderm; HB, Hindbrain; ASE, Anterior surface ectoderm; GE, Gut + Endoderm; FB, Forebrain; PIM, Posterior + Intermediate mesoderm; MB, Midbrain; HM, Head mesoderm; CM, Cardiopharyngeal mesoderm; PSE, Posterior surface ectoderm; H, Heart; SM, Somitic mesoderm; E, Endothelium; SC, Spinal cord; A, Allantois; NC, Neural crest; EE, Extraembryonic endoderm; VNT, Ventral neural tube; B, Blood; N, Notochord.

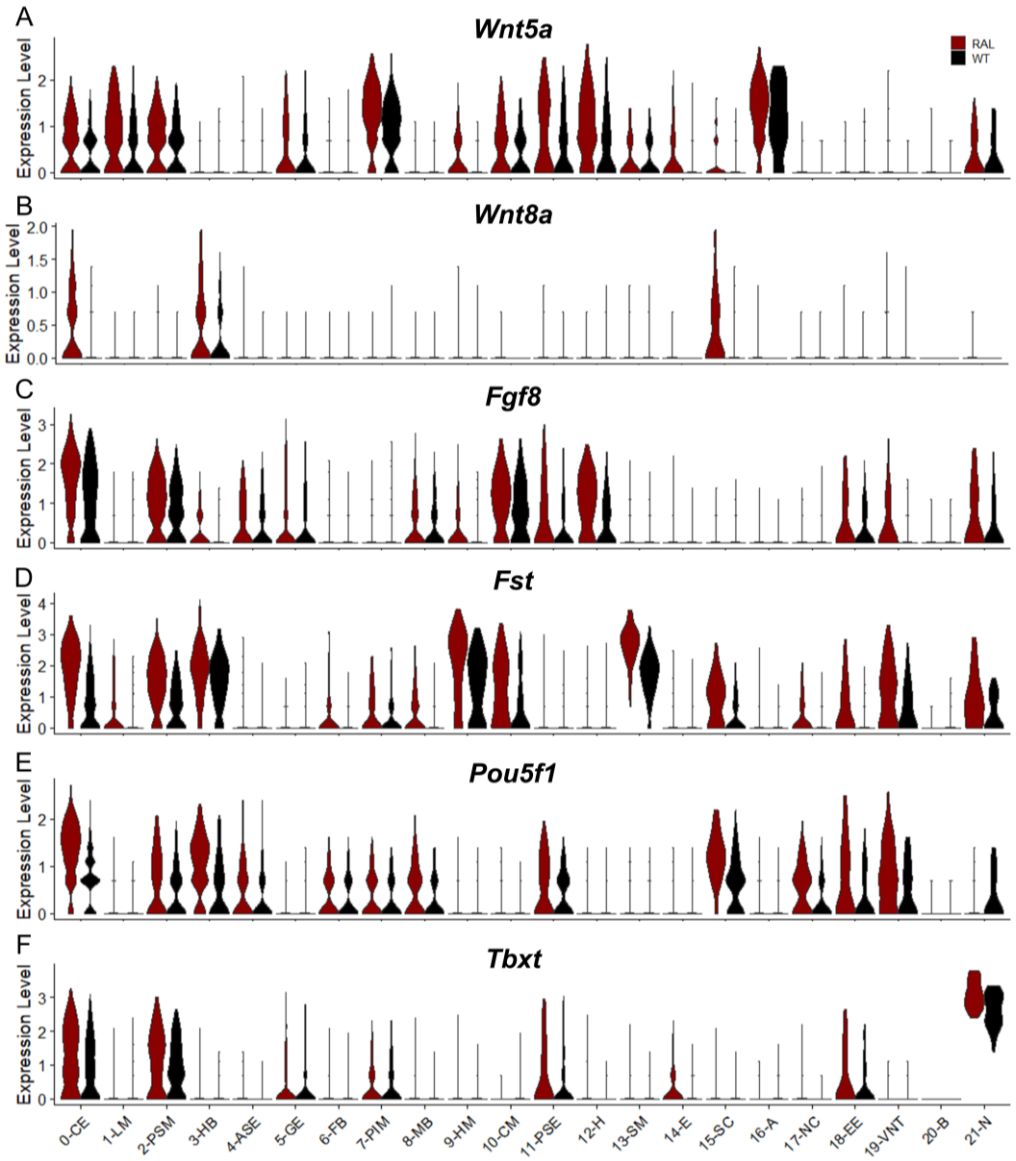


Figure 3.14 – Impact of the absence of RA signaling in the expression of key developmental genes at E8.25. (A-F) Expression levels of *Wnt5a* (A), *Wnt8a* (B), *Fgf8* (C), *Fst* (D), *Pou5f1* (E) and *Tbxt* (F) in the integrated wild type (black) and *Raldh2*$^{-/-}$ (red) datasets (as identified in Figure 3.12A). Abbreviations: CE, Caudal epiblast; LM, Lateral mesoderm; PSM, Presomitic mesoderm; HB, Hindbrain; ASE, Anterior surface ectoderm; GE, Gut + Endoderm; FB, Forebrain; PIM, Posterior + Intermediate mesoderm; MB, Midbrain; HM, Head mesoderm; CM, Cardiopharyngeal mesoderm; PSE, Posterior surface ectoderm; H, Heart; SM, Somitic mesoderm; E, Endothelium; SC, Spinal cord; A, Allantois; NC, Neural crest; EE, Extraembryonic endoderm; VNT, Ventral neural tube; B, Blood; N, Notochord.

Conclusion

In the present study, we evaluated the impact of the absence of RA signaling in both chromatin accessibility and transcriptomic profiles during embryo development. Despite the reported mechanisms for RA gene regulation relying on modulation of chromatin (Hong et al., 2018; Kumar et al., 2016; Kumar and Duester, 2014), we observed very few statistically significant changes in chromatin accessibility in the caudal epiblast of *Raldh2*$^{-/-}$ embryos at E8.5. Although RAREs have been identified near both *Nr2f2* and *Hoxa4* (Behringer et al., 1993; Berenguer et al., 2020; Packer et al., 1998), our analysis uncovered enhancer regions regulating these genes from hundreds of kilobases away, suggesting that RA signaling can have multiple mechanisms to regulate gene expression. Our functionality assays also revealed the existence of regulatory interactions, both activating and inhibiting, between enhancer regions to fine tune the expression pattern of the target gene (Bothma et al., 2015; El-Sherif and Levine, 2016; Perry et al., 2011). In addition, enhancer regions modulating the same gene were found to be regulated by distinct sets of TFs, which allow maintenance of gene expression patterns regardless of fluctuations in TF levels (Waymack et al., 2020). Overall, these results demonstrate the high complexity behind the mechanisms involved in enhancer regulation.

Single cell transcriptome profiles revealed that the absence of RA activity had a marked impact on the cellular dynamics of the caudal epiblast, in particular the disruption of progenitor cell differentiation to neural fates, consistent with previous reports (Cunningham et al., 2015a). We also identified *Fat3* as a potential driver gene of the inferred trajectories in the caudal epiblast. Although we found no significant changes of *Fat3* expression at E8.25, by E8.5 we and others (Berenguer et al., 2020) describe a downregulation in *Raldh2* mutants. These observations confirm the expression dynamics predicted by scVelo and suggest that *Fat3* potential regulatory role is compromised in *Raldh2* mutants. RA signaling has been associated with several processes

throughout development (Busada et al., 2015; Hernandez et al., 2007; Matt et al., 2005; Niederreither et al., 2001, 1999; Zhao et al., 2009). Our approach provided insight on the early RA activity at the whole embryo level, suggesting that the first regulatory defects are the extended domains of expression due to the lack of RA-dependent repression of the target genes.

Methods

Mice and embryos

Mouse embryos were recovered by cesarean section at different developmental stages and processed according to the standard methods for the distinct analysis described below. All experiments conducted on animals followed the Portuguese (Portaria 1005/92) and European (Directive 2010/63/EU) legislations concerning housing, husbandry, and welfare. The project was reviewed and approved by the Ethics Committee of *Instituto Gulbenkian de Ciência* and by the Portuguese National Entity, *Direcção Geral de Alimentação e Veterinária* (license reference 014308).

Raldh2 mutant mice were generated by introducing in frame stop codons into the second exon of the gene (Fig 3.15). A sgRNA targeting the sequence AATGGCAGAACTCAGAGAGT was generated by in vitro transcription. Briefly, oligonucleotides Raldh2-gRNA-1 and Raldh2-gRNA-2 (Table 3.3) were annealed and cloned into the BssI sites of plasmid pgRNA-basic (Casaca et al., 2016). The sgRNA was transcribed from the resulting plasmid with the MEGAshortscript T7 Kit (Life Technologies) and purified with the MEGAclear Kit (Life Technologies). Cas9 mRNA was produced by in vitro transcription from the pT7-Cas9 plasmid (Casaca et al., 2016) using the mMESSAGE mMACHINE T7 Ultra Kit (Life Technologies) and purified with the MEGAclear Kit (Life Technologies). The replacement ssDNA oligonucleotide containing three in frame stop codons followed by an EcoRI site (Raldh2-3X-Stop) (Table 3.3) was purchased from IDT. A mixture of 10 ng/μl of Cas9 mRNA, 10 ng/μl of the gRNA and 10 ng/μl of the Raldh2-3X-Stop oligonucleotide was injected into

the pronuclei of fertilized oocytes of the FVB/J background, using standard procedures (Hogan et al., 1994). The mutant allele was detected by PCR using primers Raldh2-MUT-Fw and Raldh2-MUT-Rv (Table 3.4). Targeting was confirmed by direct sequencing (Fig 3.15).

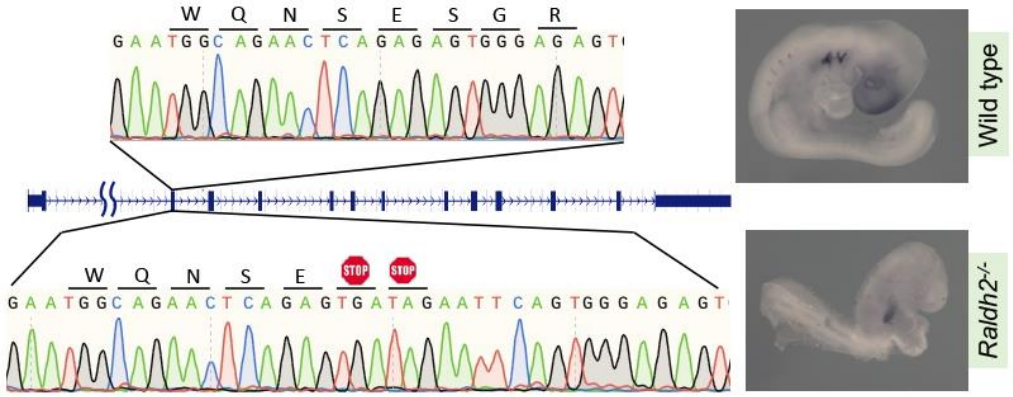


Figure 3.15 - Sequencing profiles of wild type and *Raldh2*$^{-/-}$ mutants, generated by introducing in frame stop codons in the second exon. Whole-mount in situ hybridization of wild type and *Raldh2*$^{-/-}$ mutant embryos at E9.5 using a probe for *Fgf4*.

Table 3.3 - gRNA and ssDNA used for CRISPR/Cas9

Raldh2	Raldh-gRNA-1	AGGGAATGGCAGAACTCAGAGAGT
	Raldh-gRNA-2	AAACACTCTCTGAGTTCTGCCATT
	Raldh2-3X-Stop	CATATCCCATTTTCTTGTGTCCCTTCTGTAGATCTTTATTAACAATGA ATGGCAGAACTCAGAGTGATAGAATTCAGTGGGAGAGTGTTCCCTG TCTGTAATCCAGCCACAGGAGAGCAAGTGTGTGAAGTTCAAGAAG
CR3	gRNA_CR3_1	TCTTTCGGTCGTTCCCAGAG
	gRNA_CR3_2	GATACAACCGTCTTCTAGCT
	ssDNA	CCATCGGGGCGGGTGAGACCTCTCAGCACACCCTCTGTCCCCT TCTTTCGGTCGTTCCCAGCTAGGGAACCAGGGCAAAGTTGGC CTGGGTGGGATGGTTCTAAGGGTGCAGGGTGAACA
CR3a	gRNA_CR3a_1	ATTGGAGGTGCACTGGGTGA
	gRNA_CR3_2	GATACAACCGTCTTCTAGCT
	ssDNA	GGATGCTGGTGTGTATGCTTGTATGTGCCTTTGGAGTCAGGGT ATTGGAGGTGCACTGGGGCTAGGGAACCAGGGCAAAGTTGG CCTGGGTGGGATGGTTCTAAGGGTGCAGGGTGAACA
CR3b	gRNA_CR3_1	TCTTTCGGTCGTTCCCAGAG
	gRNA_CR3b_2	TCTCCTGGGCATTATCTGCC

	ssDNA	CCATCGGGGCGGGTGAGACCTCTCAGCACACCCTCTGTCCCCT TCTTTCGGTCGTTCCCAGCCAGGTTCACCCCATTTCTTTTTATAA TCTTACTACATATTTAAAGGAGTCCCTTGCCT
CR4	gRNA_CR4_1	GAAACCTTCAGAGGCAGTAG
	gRNA_CR4_2	AGAGCTAAATCAGATGTCAC
	ssDNA	TCCGTCAATCAACAATAAAGTATTATTCTACCGTCACAAAGGGGA AACCTTCAGAGGCAG ACATCTGATTTAGCTCTTCTCCCCGAGTG ACATTCACTGCTCCTTTTCTTCACATACTTA

Table 3.4 - Primers used for genotyping.

Raldh2	MUT	Fw	GTTTTCTGATCTCCCAGATCTC
		Rv	TCTCCCACTGAATTCTATCAC
	WT	Fw	GTTTTCTGATCTCCCAGATCTC
		Rv	AACACTCTCCCACTCTCTGAG
CR3	KO	Fw	GAGCCACACTGATTTCAGAGG
		Rv	TCATCCATACCCTCCAGCTAC
	WT	Fw	GAGCCACACTGATTTCAGAGG
		Rv	AGACGTTACAGTAACGTGCTC
CR3a	KO	Fw	TGAATTGACGTGAGAGGAAGG
		Rv	TCATCCATACCCTCCAGCTAC
	WT	Fw	TGAATTGACGTGAGAGGAAGG
		Rv	GGCTGATGTGAAGCATTGCAG
CR3b	KO	Fw	GAGCCACACTGATTTCAGAGG
		Rv	TTATCACAGACTGTGACCAAC
	WT	Fw	GAGCCACACTGATTTCAGAGG
		Rv	AGACGTTACAGTAACGTGCTC
CR4	KO	Fw	CTGTGATTGAACATGGTAGGC
		Rv	TGTGACGATAAGGTCCAGTTG
	WT	Fw	CTGTGATTGAACATGGTAGGC
		Rv	GATGGAGTACACTGAGTAGTG

CR3 and CR4 deletion mice were generated by CRISPR/Cas9 (Wang et al., 2013) on the FVB/J background, as previously described (Tekko et al., 2022). We used two gRNAs targeting the border of the sequence to be deleted and one ssDNA oligo bridging the two sides of the deletion (Table 3.3) to increase the edition efficiency. In these cases, each gRNA was generated by

annealing the relevant Alt-R-CRISPR-Cas9 crRNA (targeting sequences in Table 3.3) with the Alt-R-CRISPR-Cas9 tracrRNA (all purchased from IDT). 1μ M of each gRNA was incubated with 100 ng/ml of theCas9 protein and 10 ng/ml of the replacement DNA and microinjected into the pronucleus of fertilized mouse oocytes. Identification of deletion mutants was performed by PCR using the oligonucleotides specified in Table 3.4. Positive founders were crossed with wild type mice to generate F1 heterozygous mice that were then used to build the mutant lines. Homozygous mutants were then generated by heterozygous crosses. Mice and embryos were genotyped by PCR using primers specified in Table 3.4.

ATAC-seq

Posterior epiblasts of wild type and *Raldh2*$^{-/-}$ E8.5 mouse embryos were collected to 500 μl of cold M2 (Sigma, M7167), spun down to remove supernatant and incubated with 500 μl of Accutase (Sigma, A6964) for 30 min at 37ºC, with shaking at 600 rpm, to dissociate the tissue into single cells. ATAC-seq was performed as previously described (Corces et al., 2017), using two separate biological replicates for each condition. The amplified libraries were double-step size selected (0.5x followed by 1x) using SPRIselect (Beckman Coulter, B23317) according to manufacturer's instructions. Pooled ATAC-seq libraries were sequenced on a NextSeq500 (Illumina) at 50M paired-end 75 base reads per sample.

ATAC-seq Data Analysis

Fastq files were processed with GUAVA v1, following the recommended guidelines (Divate and Cheung, 2018). GUAVA enables pre-processing of raw sequencing reads, mapping of reads to a reference genome, peak calling and annotation, as well as differential analysis between samples. ATAC-seq data is available in the GEO accession database under accession number GSE220245. All genome browser tracks were captured using Integrative

Genomics Viewer (Robinson et al., 2011). Phylogenetic conservation data for multiple alignments of 59 vertebrate genomes to the mouse genome (mm10.60way.phyloP60way) was obtained from phyloP directory (Pollard et al., 2010) and loaded into IGV. Hi-C data was obtained from 3D genome browser (Wang et al., 2018) using the 'mm10 mESC Bonev_2017-raw' dataset at 40kb resolution. To visualize our candidate regions within the context of this dataset we loaded a BED file with the coordinates of our candidate regions to UCSC Genome Browser (Kent et al., 2002) and loaded this session into the 3D genome browser (Wang et al., 2018).

ATAC-seq data was analyzed for TF footprints using HINT (Li et al., 2019). Replicates were merged to increase read depth and processed with "rgt-hint footprinting" command. Footprint motifs were matched to HOCOMOCO database (Kulakovskiy et al., 2018) with "rgt-motifanalysis matching" and then further assessed for differential motif occupancy with the "rgt-hint differential" command.

β-Galactosidase Transgenics

For reporter analyses, candidate regions identified by ATAC-seq data were amplified by PCR from mouse genomic DNA (primers provided below, Table 3.5) and cloned upstream of a cassette containing the adenovirus 2 minimal late promoter, the β-galactosidase cDNA, and the polyadenylation signal from SV40 (Jurberg et al., 2013). Transgenic mice were produced by pronuclear injection (Hogan et al., 1994). The β-galactosidase staining was performed as previously described (Jurberg et al., 2013). At least 3 transgenic embryos displaying consistent reporter expression patterns were observed for each candidate.

Table 3.5 - Primers used to amplify candidate regions for β-Galactosidase assays.

CR3	Fw	GACTCGAGGTGTCAGACCTGTGTAAATGC
	Rv	ACCTGCAGGGAGGAAATGTTGTTGTTTGG
CR3a	Fw	TTACTCGAGGCTGTCTACAGTGACTCTGTG

	Rv		ACCTGCAGGGAGGAAATGTTGTTGTTTGG
CR3b	Fw		GACTCGAGGTGTCAGACCTGTGTAAATGC
	Rv		ATCTGCAGGTAAGGAGCAGACTTCACGTC
CR3a$^{\Delta Msgn1}$	Up	Fw	TTACTCGAGGCTGTCTACAGTGACTCTGTG
		Rv	GCCGAATTCTACTAGTTTATGGGGCTGATG
	Down	Fw	CGCGAATTCGACACTTGAAAGTACCAGTTC
		Rv	ACCTGCAGGGAGGAAATGTTGTTGTTTGG
CR3a$^{\Delta Hox}$	Up	Fw	TTACTCGAGGCTGTCTACAGTGACTCTGTG
		Rv	GCAGAATTCGGCTGATGTGAAGCATTGCAG
	Down	Fw	GCCGAATTCAGTCAGAACAAAAGGTCTGAC
		Rv	ACCTGCAGGGAGGAAATGTTGTTGTTTGG
CR3$^{\Delta Hox}$	Up	Fw	GACTCGAGGTGTCAGACCTGTGTAAATGC
		Rv	GCAGAATTCGGCTGATGTGAAGCATTGCAG
	Down	Fw	GCCGAATTCAGTCAGAACAAAAGGTCTGAC
		Rv	ACCTGCAGGGAGGAAATGTTGTTGTTTGG
CR3b$^{\Delta Smad1+Sp5}$	Up	Fw	GACTCGAGGTGTCAGACCTGTGTAAATGC
		Rv	ATTCTAGAGTGGCTTCTGCTCCAGAGCTC
	Down	Fw	CGTCTAGATTCTAAGAGACTCAGTGGCTC
		Rv	ATCTGCAGGTAAGGAGCAGACTTCACGTC
CR4	Fw		GCACTCGAGAGTGGTTATAAGCTTGCCTTG
	Rv		GTCTGCAGAAATCAGATGTCACTGGAGTG
CR4$^{\Delta Rfx6+Pax3}$	Up	Fw	GCACTCGAGAGTGGTTATAAGCTTGCCTTG
		Rv	GGTCTAGAACAACCCATAGTCCTCTACTG
	Down	Fw	AATCTAGACGCTGGCTTGGGGCAAGAAGC
		Rv	GTCTGCAGAAATCAGATGTCACTGGAGTG
CR4$^{\Delta Vdr}$	Up	Fw	GCACTCGAGAGTGGTTATAAGCTTGCCTTG
		Rv	CCTCTAGACCAGCGGTCAGCAACAAATAATTC
	Down	Fw	GCTCTAGATTCCTGATGAGCACCATCACC
		Rv	GTCTGCAGAAATCAGATGTCACTGGAGTG
Peak_10773	Fw		GTACTCGAGGCCAGGCTAAAGAGTATGTTG
	Rv		AACTGCAGTGTATGCTGGCTTGGCTGCTG
Peak_11325	Fw		CGACTCGAGGTGAGTGTATTTCACAGAAGC
	Rv		CGCAGATCTCCCAATTCTGTGTGAAATTGG
Peak_14942	Fw		GTCTCGAGAGGCTGGATGACTAATGGTTAGC
	Rv		AACTGCAGCTTCTGCGTTGTCTGTGAGTG

Peak_35261	Fw	GTCTCGAGTGACACGGAGTATTTCATGTC
	Rv	AGCTGCAGCTCAAACTTCTTCCTAACCTG
Peak_59634	Fw	CATCTCGAGTAAAGCAGAGTCTGAAGGCAG
	Rv	GTCAGATCTTCAAAGCTTTGCCAGAGCCAG
Peak_60984	Fw	GCGCTCGAGTCCACACTGAACTATTTGTTC
	Rv	CACTGCAGTCTTCTCTCACACATTCCAAC
Peak_67886	Fw	GTACTCGAGTGCTGGGATCTCAAAAGACAG
	Rv	AGCTGCAGGCTACTGTGTATCTGAACATG
Peak_123811	Fw	CGACTCGAGAGTTCCACTAAAAGGCACCATC
	Rv	TTCTGCAGGATTACAGCAATTGGCACCAC
Peak_143897	Fw	ACGCTCGAGAGGACCACTTAACAATTCTGC
	Rv	TCCTGCAGTAACCTACTACACCAAATGGCC
Peak_152170	Fw	CGGCTCGAGGCATACATTAGTCTTTCTTCACC
	Rv	TGCAGATCTCCTTGTGCATTCAGAGTGCAG

Whole-mount *in situ* hybridization

Whole-mount in situ hybridization was performed as previously described (Aires et al., 2019) using digoxigenin-labeled RNA antisense probes. At least 3 embryos were stained per probe and genotype, showing highly reproducible patterns. RNA probes were prepared by amplifying a cDNA fragment by RT-PCR (primers provided below, Table 3.6) from total RNA isolated from E9.5 embryos with TRI Reagent (Sigma, 93289), according to the manufacturer's protocol, and cloning it into pKS-bluescript. The *Hoxa6* probe was obtained from a cDNA image clone (MGC:155423, IMAGE:8733856).

Table 3.6 - Primers used to amplify in situ probes.

Nr2f2	Fw	ACGAATTCTGCATGCAGCCTAACAACATC
	Rv	ATGGATCCATTGCTCTATGACTGAGGAGG
Hoxa4	Fw	GCGAATTCCACTTTAACCGCTACCTGACC
	Rv	TTGGATCCTTCCACTAAGCAGTAACGAGG
Hoxa5	Fw	GCGAATTCATCATAGTTCCGTGAGCGAAC
	Rv	TTGGATCCTCTTTCTCCAGCTCCAGGGTC
Hoxa7	Fw	GGATCGATGCTTATACAATGTCAACAGCC

	Rv	AAGGATCCCTGGCTCTCATCTTTATGCTC

RT-qPCR

Total RNA was extracted from whole wild type and *CR4*$^{\Delta/\Delta}$ embryos at E9.5 using TRI Reagent. 1 μg of RNA was used for reverse transcription into complementary DNA (cDNA) using NZY Reverse Transcriptase enzyme (NZYTech, MB124) and random hexamer mix (NZYTech, MB12901) following the manufacturer's protocol. Real-time qPCR was performed in a QuantStudio 7 Flex real-time PCR system (Thermo Fisher) using iQ SYBR Green Supermix (Bio-rad, 1708880) according to manufacturer's instructions. Primers used are listed in Table 3.7. Quantification was determined using the standard curve method, and expression levels normalized to *β-Actin*. Statistical significance was assessed using unpaired t-test.

Table 3.7 - Primers used in RT-qPCR and ChIP-qPCR.

Hoxa4	Fw	CCTGGATGAAGAAGATCCACG
	Rv	TTCCTTCTCCAGTTCCAAGAC
β-Actin	Fw	TCTGGTGGTACCACCATGTAC
	Rv	TACTTGCGCTCAGGAGGAGC
CR4-Vdr	Fw	TTCCACCAGCCACAAACTTAG
	Rv	TGGAGTACACTGAGTAGTGGC

ChIP-qPCR

Posterior caudal epiblasts of wild type embryos were collected at E8.5 for ChIP as previously described (Schmidl et al., 2015). Immunoprecipitation was performed using rabbit monoclonal anti-Vitamin D Receptor antibody (Abcam, Ab109234), and rabbit polyclonal IgG antibody (Cell Signaling, 2729) was used as a negative control. Enrichment was measured by real-time qPCR in a QuantStudio 7 Flex real-time PCR system (Thermo Fisher) using iQ SYBR Green Supermix (Bio-rad, 1708880) according to manufacturer's instructions. Primers used are listed in Table 3.7. Quantification was determined using the

relative quantification method, enrichment was normalized to IgG. Statistical significance was assessed using unpaired t-test.

RNA-sequencing analysis

Posterior epiblasts of wild type and *Raldh2*$^{-/-}$ E8.5 mouse embryos were dissected and snap frozen. Total RNA was isolated from pooled samples with TRI Reagent following the manufacturer's protocol. RNA samples were then resuspended in RNase-free water. RNA concentration and purity were determined on an AATI Fragment Analyzer (Agilent). RNA-seq from wild type and *Raldh2*$^{-/-}$ tissues was performed using two separate biological replicates. Libraries were prepared from total RNA using the SMART-Seq2 protocol (Picelli et al., 2014). Sequencing was performed on Illumina NextSeq500, generating >25M single-end 75 base reads per sample. Reads were aligned to the reference mouse genome (mm10) using STAR (Dobin et al., 2013). Read count normalization and differential expression between samples was analyzed using DESeq2 (Love et al., 2014). Gene ontology enrichment analysis was performed using PANTHER (Ashburner et al., 2000; Mi et al., 2019), by selecting for biological processes using Fisher's Exact test and False Discovery Rate. No background gene list was used. Gene ontology results presented are ranked by Fold Enrichment. RNA-seq data is available in the GEO accession database under the accession number GSE220246.

Single-cell RNA sequencing

Embryos from *Raldh2* heterozygous crossings were collected at E8.25 and dissected in M2 media (Sigma, M7167). Since at this developmental stage homozygous mutants are not visually distinguishable, yolk sacs were collected for genotyping while the dissected embryos were maintained in M2 media on ice. Samples were genotyped using Speedy Supreme NZYTaq DNA polymerase (NZYtech, MB360), following the manufacturer's instructions. Whole wild type and *Raldh2*$^{-/-}$ embryos were pooled for one sample of each genotype. Pooled samples were dissociated into single cell suspension by

incubating the embryos with TrypLE Express (Life Technologies, 12604-013) at 37ºC for 7 min followed by mechanical dissociation by pipetting and quenched with M2 supplemented with 10% FBS (GE Healthcare, HYCLSH30070.03). Single cell suspensions were centrifuged at 300g for 5min, resuspended in PBS with 0,04% UltraPure BSA (Ambion, AM2616) and filtered through a 40µm cell strainer (PluriSelect, SKU 43-10040-40). Cell counts were determined with a haemocytometer. Single cell libraries were generated with Chromium Next GEM Single Cell 3' Kit (10x Genomics, v3.1), according to manufacturer's instructions, using 10 000 cells as input. Libraries were sequenced on Illumina NextSeq500 with High Output v2.5 (75 cycles) kit (Illumina, 20024906).

Raw fasta files were processed with 10x Genomics Cell Ranger pipeline (10x Genomics, v3.0), reads were mapped to the mouse genome (mm10) and counted using UMI barcodes to generate count matrices. Cell quality control, data normalization, clustering and differential analysis was performed using Seurat V4.0 package (Stuart et al., 2019), following developers' recommendations. Cells with less than 2000 expressed genes, as well as with more than 5% of mitochondrial genes were excluded from the analysis. RNA velocity was computed using Velocyto (La Manno et al., 2018) and scVelo (Bergen et al., 2020), according to the developer's instructions.

Acknowledgments

We thank the IGC animal house and Genomics Facilities for their expert services, advice and assistance, as well as António Sousa from the Bioinformatics Unit for performing RNA velocity analysis. We also acknowledge all members of the Mallo laboratory for helpful discussions and comments throughout the course of the project.

This work was supported by the Fundação para a Ciência e a Tecnologia (LISBOA-01-0145-FEDER-030254) to M.M., by PhD fellowships from the Fundação para a Ciência e a Tecnologia (PD/BD/138240/2018 &

COVID/BD152639/2022) to P.D. and (PD/BD/128426/2017) to A.D., and by the research infrastructure Congento (LISBOA-01-0145-FEDER-022170).

Chapter 4

General Discussion

The vertebrate main body axis is generated sequentially from head to tail, first by the activity of the primitive streak and then of the tail bud. Despite the diversity of body shapes and sizes present in this clade, they all share a basic organization of their body plan, distributed into head, neck, trunk and tail structures. Head and trunk development are distinct not only in terms of cellular dynamics and types of tissues produced, but also in the gene regulatory networks controlling their development. In chapter 2 we described a wide variety of molecular and genomic changes associated with the head to trunk transition and further focused on relevant changes observed in WNT signaling. *Wnt5a* and *Wnt11*, non-canonical Wnt ligands required for axial extension through the trunk and tail (Andre et al., 2015; Yamaguchi et al., 1999), were found upregulated already at E8.5, suggesting an increased reliance on non-canonical WNT signaling in the posterior epiblast upon head to trunk transition. This was also accompanied by significant changes in G-protein-dependent signaling that may play a functional role in the activation of the non-canonical WNT pathway.

We also found that the effect of Porcn on WNT signaling might be different during head and trunk development. Porcn introduces a palmitoleoyl moiety to Wnt proteins (Rios-Esteves et al., 2014; Takada et al., 2006) and experiments in cultured cells indicated that its absence results in Wnt protein accumulation in the endoplasmic reticulum. *Porcn* inactivation in mice produced development arrest during gastrulation (Biechele et al., 2011; Takada et al., 2006), indicating an essential role of this enzyme for Wnt3 activity. It has been assumed that Porcn is also required for WNT signaling at later developmental stages of axial development, although it had never been directly evaluated. The finding that Porcn is downregulated at E8.5, an observation that fits with published expression analyses during embryonic development (Biechele et al., 2011), led us to analyze the effect of blocking Porcn activity on axial extension. Our ex-vivo embryo culture assays in the presence of the Porcn inhibitor IWP-01, revealed that WNT signaling during axial extension might include both Porcn-dependent and -independent activities. Significantly, the phenotype of IWP-01

treated embryos did not include the formation of ectopic neural tubes at the expense of paraxial mesoderm, a hallmark of the absence of *Wnt3a* in mouse embryos (Yoshikawa et al., 1997). This indicates that, at least the known Wnt3a activity as a coordinator of NMC fate is independent of Porcn activity. These findings are consistent with previous reports showing that Wnt3a can activate WNT signaling independently of Porcn in cell culture assays (Rao et al., 2019). Interestingly, IWP-01-treated embryos presented a phenotype reminiscent of *Wnt5a* mutants (Yamaguchi et al., 1999), indicating that its function during axial extension requires posttranslational modification by Porcn. In addition, this phenotype is also consistent with an increased reliance on non-canonical WNT signaling during axial extension. Wnt5a has been described to inhibit canonical WNT signaling during secondary heart field development by modulating the levels of β-catenin (Bisson et al., 2015; Cohen et al., 2012). It would be interesting to determine whether a similar mechanism could operate during axial extension, to establish a proper functional balance between canonical and non-canonical signaling. Such balance could be required both to regulate the promotion of axial extension and to achieve a proper equilibrium between the production of neural tube and paraxial mesoderm from the NMCs.

Key regulators of trunk development like *Cdx2, Tbxt* and *Wnt3a* only seem to become functionally relevant during the head to trunk transition, despite already being expressed at earlier developmental stages (Amin et al., 2016; Herrmann et al., 1990; Takada et al., 1994). Inactivation of these genes originates truncated embryos at the level of this transition, while the formation of anterior structures is not affected (Amin et al., 2016; Herrmann et al., 1990; Takada et al., 1994). This suggests that the head to trunk transition may involve a change in the capacity of cells to respond to specific gene networks. Consistent with this hypothesis, we observed extensive changes in chromatin accessibility, with most of the differentially accessible regions mapping to intergenic regions, thus indicating a major switch in the regulatory elements associated with either head or trunk formation. Analysis of TF binding dynamics revealed that the binding activity of Cdx2, Cdx1, and several posterior Hox

proteins only kickstarts when trunk formation is initiated. This switch in binding activity demonstrates how changes in chromatin accessibility can determine which regulatory networks become active or inactive at a particular developmental stage by controlling the expression of genes playing relevant roles in those networks. Cdx2 has been shown to activate genes involved in FGF and WNT signaling pathways by binding to regulatory elements controlling their activity (Amin et al., 2016). One of the genomic regions bound by Cdx2 is located upstream of *Wnt5a* (Amin et al., 2016) and overlaps with the CR1 element we described herein. Together, this suggests that the change in chromatin state during the head to trunk transition may be required to enable Cdx2 binding to the CR1 region, which in turn will promote proper timing of *Wnt5a* expression.

RA signaling is also required for proper regulation of the head to trunk transition, as inactivation of *Raldh2* results in developmental arrest at this stage (Niederreither et al., 1999). Despite our initial hypothesis of RA acting as a regulator of the chromatin changes observed during this transition, absence of RA activity had a minor effect in chromatin accessibility in the caudal epiblast at E8.5. This was somewhat unexpected, considering previous reports showing that RA signaling has a marked impact on chromatin accessibility (Kiani et al., 2022; Li et al., 2018; López-Pérez et al., 2021; Schleif et al., 2023). However, in all of these studies chromatin dynamics were assessed following addition of RA, which may generate non-physiological effects, as the described RA effect on chromatin accessibility was dose dependent (Kiani et al., 2022).

Although *Raldh2* mutant embryos still do not display a visible phenotype at E8.25, single cell transcriptome analysis of the whole embryo revealed that the cellular dynamics of the caudal epiblast was severely disturbed. This was also accompanied by a reduction of the spinal cord cell population size in *Raldh2* mutants. These findings are consistent with the requirement of RA signaling for proper balance of NMC differentiation into mesodermal and neural fates (Cunningham et al., 2015a) as well as a possible role for RA in cell proliferation in the spinal cord (England et al., 2011; Wilson et al., 2003). Analysis of the

genes with the highest differential RNA velocity identified *Fat3* as a potential regulator of this process, which is supported by its requirement for neural progenitor proliferation (Seo et al., 2022). While at E8.25 *Fat3* expression levels are similar, by E8.5 *Fat3* was downregulated in *Raldh2* mutants, thus confirming the expression dynamics estimated with scVelo. Therefore, the reduced expression of *Fat3* may be responsible for the decrease in neural progenitors in *Raldh2* mutants (Cunningham et al., 2015a).

We also observed a global upregulation in *Raldh2* mutants of several genes encoding for components of the electron transport chain. Further studies will be required to determine whether this upregulation is also accompanied by increased activity of the electron transport chain in *Raldh2*$^{-/-}$. A recent study has demonstrated that increased electron transport chain activity can accelerate the segmentation clock, by regulating the NAD^+/NADH ratio (Diaz-Cuadros et al., 2023). In turn, a faster segmentation clock rate leads to smaller somites (Gomez et al., 2008). Therefore, this mechanism could potentially account for the smaller somites observed in *Raldh2* mutants. Interestingly, at E8.25 the majority of the differentially expressed genes between wild type and *Raldh2*$^{-/-}$ embryos were upregulated in the mutant dataset, likely due to the lack of repression by RA signaling. Some of these genes, like *Fgf8, Pou5f1* and *Fst*, had already been shown to be repressed by RA (Cunningham et al., 2016; Gupta et al., 2008; Kumar and Duester, 2014).

In this work, we also characterized the functional relevance of several potential enhancers identified in the chromatin accessibility assays. Generation of CR deletion mutants revealed no obvious phenotypic defects for any of the tested elements. Even though in specific cases single enhancer deletion can lead to significant phenotypes (Banerji et al., 1981; Lettice et al., 2003; Yanagisawa et al., 2003), there is growing evidence for the existence of widespread enhancer redundancy (Ahituv et al., 2007; Antosova et al., 2016; Cannavò et al., 2016; Cunningham et al., 2018; Nolte et al., 2014; Osterwalder et al., 2018). Indeed, approximately 340,000 candidate cis-regulatory elements have been identified in the mouse genome, which vastly surpasses the number

of annotated genes (around 24,000 with protein sequence data) (The ENCODE Project Consortium et al., 2020). In keeping with this global estimative, we also found a high disparity between the number of genomic regions presenting significant changes in accessibility and the differentially expressed genes associated with the head to trunk transition. These numbers point towards a complex regulatory mechanism where each gene may be controlled by multiple enhancers. The presence of redundancy between enhancers provides phenotypic robustness by protecting against genetic and environmental perturbations (Antosova et al., 2016; Frankel et al., 2010; Osterwalder et al., 2018; Perry et al., 2010).

Importantly, the redundancy that enhancers display most often does not include sequence homology (Barth et al., 2020). Our observations with CR3a and CR3b can illustrate this phenomenon, as these genomic regions could induce partially overlapping patterns of expression, but their regulation was dependent on distinct TFs. Indeed, redundant enhancer pairs have been shown to drive more robust gene expression patterns than a mere enhancer duplication (Waymack et al., 2020; Wunderlich et al., 2015). This is due to the regulation of each enhancer relying on a different set of TFs, which would help buffering the effects of possible fluctuations in the levels of their regulatory TFs (Waymack et al., 2020; Wunderlich et al., 2015). We also observed positive and negative enhancer interactions between CR3a and CR3b that were required to reproduce the expression pattern of the target gene, *Nr2f2*, in a way that each individual enhancer could not. Our observations with the CR3 enhancer are consistent with previous reports on the role of enhancer interaction in producing precise boundaries in the expression domains of developing *Drosophila* embryos (Bothma et al., 2015; El-Sherif and Levine, 2016; Perry et al., 2011). As such, redundant enhancers can present individual functional roles in what concerns time- or cell-specific regulation, which may favor their prevalence in the genome and explain why these elements are maintained throughout evolution.

Overall, by exploring both transcriptome and chromatin accessibility profiles of mouse embryos we demonstrated that the head to trunk transition is associated with extensive changes in chromatin accessibility that might play a fundamental role in the establishment of proper time- and cell- specific activation of the relevant gene regulatory networks. Despite the importance of proper regulation of the head to trunk transition for the normal course of embryonic development, this has rarely been the focus of mechanistic analysis of developmental processes. Further efforts must be employed to fully understand the control of the rapid cellular and molecular changes occurring at this embryonic stage. In particular, it would be important to identify a possible master regulator of the head to trunk transition, somewhat similar to what has been previously described for the regulation of the trunk to tail transition (Dias et al., 2020; Jurberg et al., 2013).

References

Abb, J., Abb, H., Deinhardt, F., 1982. Effect of retinoic acid on the spontaneous and interferon-induced activity of human natural killer cells. International Journal of Cancer 30, 307–310. https://doi.org/10.1002/ijc.2910300309

Abedini, A., Sayed, C., Carter, L.E., Boerboom, D., Vanderhyden, B.C., 2020. Non-canonical WNT5a regulates Epithelial-to-Mesenchymal Transition in the mouse ovarian surface epithelium. Sci Rep 10, 9695. https://doi.org/10.1038/s41598-020-66559-9

Aberle, H., Bauer, A., Stappert, J., Kispert, A., Kemler, R., 1997. β-catenin is a target for the ubiquitin–proteasome pathway. The EMBO Journal 16, 3797–3804. https://doi.org/10.1093/emboj/16.13.3797

Abrami, L., Kunz, B., Iacovache, I., van der Goot, F.G., 2008. Palmitoylation and ubiquitination regulate exit of the Wnt signaling protein LRP6 from the endoplasmic reticulum. Proceedings of the National Academy of Sciences 105, 5384–5389. https://doi.org/10.1073/pnas.0710389105

Adamska, M., Billi, A.C., Cheek, S., Meisler, M.H., 2005. Genetic interaction between Wnt7a and Lrp6 during patterning of dorsal and posterior structures of the mouse limb. Developmental Dynamics 233, 368–372. https://doi.org/10.1002/dvdy.20437

Afzal, Z., Krumlauf, R., 2022. Transcriptional Regulation and Implications for Controlling Hox Gene Expression. Journal of Developmental Biology 10, 4. https://doi.org/10.3390/jdb10010004

Agarwal, P., Wylie, J.N., Galceran, J., Arkhitko, O., Li, C., Deng, C., Grosschedl, R., Bruneau, B.G., 2003. Tbx5 is essential for forelimb bud initiation following patterning of the limb field in the mouse embryo. Development 130, 623–633. https://doi.org/10.1242/dev.00191

Agrawal, P., Blinka, S., Pulakanti, K., Reimer, M.H., Stelloh, C., Meyer, A.E., Rao, S., 2021. Genome editing demonstrates that the -5 kb Nanog enhancer regulates Nanog expression by modulating RNAPII initiation and/or recruitment. J Biol Chem 296, 100189. https://doi.org/10.1074/jbc.RA120.015152

Ahituv, N., Zhu, Y., Visel, A., Holt, A., Afzal, V., Pennacchio, L.A., Rubin, E.M., 2007. Deletion of Ultraconserved Elements Yields Viable Mice. PLOS Biology 5, e234. https://doi.org/10.1371/journal.pbio.0050234

Ahn, J.Y., Lee, J.T., 2010. Retinoic acid accelerates downregulation of the Xist repressor, Oct4, and increases the likelihood of Xist activation when Tsix is deficient. BMC Dev Biol 10, 90. https://doi.org/10.1186/1471-213X-10-90

Aiken, C.E.M., Swoboda, P.P.L., Skepper, J.N., Johnson, M.H., 2004. The direct measurement of embryogenic volume and nucleo-cytoplasmic ratio during mouse pre-implantation development. Reproduction 128, 527–535. https://doi.org/10.1530/rep.1.00281

Aires, R., de Lemos, L., Nóvoa, A., Jurberg, A.D., Mascrez, B., Duboule, D., Mallo, M., 2019. Tail Bud Progenitor Activity Relies on a Network Comprising Gdf11,

Lin28, and Hox13 Genes. Developmental Cell 48, 383-395.e8. https://doi.org/10.1016/j.devcel.2018.12.004

Aires, R., Dias, A., Mallo, M., 2018. Deconstructing the molecular mechanisms shaping the vertebrate body plan. Current Opinion in Cell Biology, Differentiation and disease 55, 81–86. https://doi.org/10.1016/j.ceb.2018.05.009

Aires, R., Jurberg, A.D., Leal, F., Nóvoa, A., Cohn, M.J., Mallo, M., 2016. Oct4 Is a Key Regulator of Vertebrate Trunk Length Diversity. Dev Cell 38, 262–274. https://doi.org/10.1016/j.devcel.2016.06.021

Allan, J., Hartman, P.G., Crane-Robinson, C., Aviles, F.X., 1980. The structure of histone H1 and its location in chromatin. Nature 288, 675–679. https://doi.org/10.1038/288675a0

Amin, S., Neijts, R., Simmini, S., van Rooijen, C., Tan, S.C., Kester, L., van Oudenaarden, A., Creyghton, M.P., Deschamps, J., 2016. Cdx and T Brachyury Co-activate Growth Signaling in the Embryonic Axial Progenitor Niche. Cell Reports 17, 3165–3177. https://doi.org/10.1016/j.celrep.2016.11.069

Andre, P., Song, H., Kim, W., Kispert, A., Yang, Y., 2015. Wnt5a and Wnt11 regulate mammalian anterior-posterior axis elongation. Development 142, 1516–1527. https://doi.org/10.1242/dev.119065

Angers, S., Thorpe, C.J., Biechele, T.L., Goldenberg, S.J., Zheng, N., MacCoss, M.J., Moon, R.T., 2006. The KLHL12–Cullin-3 ubiquitin ligase negatively regulates the Wnt–β-catenin pathway by targeting Dishevelled for degradation. Nat Cell Biol 8, 348–357. https://doi.org/10.1038/ncb1381

Antosova, B., Smolikova, J., Klimova, L., Lachova, J., Bendova, M., Kozmikova, I., Machon, O., Kozmik, Z., 2016. The Gene Regulatory Network of Lens Induction Is Wired through Meis-Dependent Shadow Enhancers of Pax6. PLOS Genetics 12, e1006441. https://doi.org/10.1371/journal.pgen.1006441

Ara, T., Nakamura, Y., Egawa, T., Sugiyama, T., Abe, K., Kishimoto, T., Matsui, Y., Nagasawa, T., 2003. Impaired colonization of the gonads by primordial germ cells in mice lacking a chemokine, stromal cell-derived factor-1 (SDF-1). Proc Natl Acad Sci U S A 100, 5319–5323. https://doi.org/10.1073/pnas.0730719100

Arnold, S.J., Hofmann, U.K., Bikoff, E.K., Robertson, E.J., 2008. Pivotal roles for eomesodermin during axis formation, epithelium-to-mesenchyme transition and endoderm specification in the mouse. Development 135, 501–511. https://doi.org/10.1242/dev.014357

Asada, H., Kawamura, Y., Maruyama, K., Kume, H., Ding, R.G., Kanbara, N., Kuzume, H., Sanbo, M., Yagi, T., Obata, K., 1997. Cleft palate and decreased brain gamma-aminobutyric acid in mice lacking the 67-kDa isoform of glutamic acid decarboxylase. Proc Natl Acad Sci U S A 94, 6496–6499. https://doi.org/10.1073/pnas.94.12.6496

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin,

G.M., Sherlock, G., 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25, 25–29. https://doi.org/10.1038/75556

Aulehla, A., Wehrle, C., Brand-Saberi, B., Kemler, R., Gossler, A., Kanzler, B., Herrmann, B.G., 2003. Wnt3a Plays a Major Role in the Segmentation Clock Controlling Somitogenesis. Developmental Cell 4, 395–406. https://doi.org/10.1016/S1534-5807(03)00055-8

Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., Lovell-Badge, R., 2003. Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev. 17, 126–140. https://doi.org/10.1101/gad.224503

Banerji, J., Rusconi, S., Schaffner, W., 1981. Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell 27, 299–308. https://doi.org/10.1016/0092-8674(81)90413-x

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., Kouzarides, T., 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature 410, 120–124. https://doi.org/10.1038/35065138

Bänziger, C., Soldini, D., Schütt, C., Zipperlen, P., Hausmann, G., Basler, K., 2006. Wntless, a Conserved Membrane Protein Dedicated to the Secretion of Wnt Proteins from Signaling Cells. Cell 125, 509–522. https://doi.org/10.1016/j.cell.2006.02.049

Bardot, E.S., Hadjantonakis, A.-K., 2020. Mouse gastrulation: Coordination of tissue patterning, specification and diversification of cell fate. Mechanisms of Development 163, 103617. https://doi.org/10.1016/j.mod.2020.103617

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., Zhao, K., 2007. High-Resolution Profiling of Histone Methylations in the Human Genome. Cell. https://doi.org/10.1016/j.cell.2007.05.009

Barth, N.K.H., Li, L., Taher, L., 2020. Independent Transposon Exaptation Is a Widespread Mechanism of Redundant Enhancer Evolution in the Mammalian Genome. Genome Biology and Evolution 12, 1–17. https://doi.org/10.1093/gbe/evaa004

Beccari, L., Yakushiji-Kaminatsui, N., Woltering, J.M., Necsulea, A., Lonfat, N., Rodríguez-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. https://doi.org/10.1101/gad.281055.116

Becchetti, A., Munaron, L., Arcangeli, A., 2013. The role of ion channels and transporters in cell proliferation and cancer. Front Physiol 4, 312. https://doi.org/10.3389/fphys.2013.00312

Bednar, J., Horowitz, R.A., Grigoryev, S.A., Carruthers, L.M., Hansen, J.C., Koster, A.J., Woodcock, C.L., 1998. Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. Proc Natl Acad Sci U S A 95, 14173–14178. https://doi.org/10.1073/pnas.95.24.14173

Behrens, J., von Kries, J.P., Kühl, M., Bruhn, L., Wedlich, D., Grosschedl, R., Birchmeier, W., 1996. Functional interaction of beta-catenin with the transcription factor LEF-1. Nature 382, 638–642. https://doi.org/10.1038/382638a0

Behringer, R.R., Crotty, D.A., Tennyson, V.M., Brinster, R.L., Palmiter, R.D., Wolgemuth, D.J., 1993. Sequences 5' of the homeobox of the Hox-1.4 gene direct tissue-specific expression of lacZ during mouse development. Development 117, 823–833. https://doi.org/10.1242/dev.117.3.823

Belo, J.A., Bouwmeester, T., Leyns, L., Kertesz, N., Gallo, M., Follettie, M., De Robertis, E.M., 1997. Cerberus-like is a secreted factor with neuralizing activity expressed in the anterior primitive endoderm of the mouse gastrula. Mechanisms of Development 68, 45–57. https://doi.org/10.1016/S0925-4773(97)00125-1

Benbrook, D., Lernhardt, E., Pfahl, M., 1988. A new retinoic acid receptor identified from a hepatocellular carcinoma. Nature 333, 669–672. https://doi.org/10.1038/333669a0

Berenguer, M., Meyer, K.F., Yin, J., Duester, G., 2020. Discovery of genes required for body axis and limb formation by global identification of retinoic acid–regulated epigenetic marks. PLoS Biol 18, e3000719. https://doi.org/10.1371/journal.pbio.3000719

Bergen, V., Lange, M., Peidli, S., Wolf, F.A., Theis, F.J., 2020. Generalizing RNA velocity to transient cell states through dynamical modeling. Nat Biotechnol 38, 1408–1414. https://doi.org/10.1038/s41587-020-0591-3

Berger, M.F., Badis, G., Gehrke, A.R., Talukder, S., Philippakis, A.A., Peña-Castillo, L., Alleyne, T.M., Mnaimneh, S., Botvinnik, O.B., Chan, E.T., Khalid, F., Zhang, W., Newburger, D., Jaeger, S.A., Morris, Q.D., Bulyk, M.L., Hughes, T.R., 2008. Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. Cell 133, 1266–1276. https://doi.org/10.1016/j.cell.2008.05.024

Berkes, C.A., Bergstrom, D.A., Penn, B.H., Seaver, K.J., Knoepfler, P.S., Tapscott, S.J., 2004. Pbx Marks Genes for Activation by MyoD Indicating a Role for a Homeodomain Protein in Establishing Myogenic Potential. Molecular Cell 14, 465–477. https://doi.org/10.1016/S1097-2765(04)00260-6

Bernheim, S., Meilhac, S.M., 2020. Mesoderm patterning by a dynamic gradient of retinoic acid signalling. Philosophical Transactions of the Royal Society B: Biological Sciences 375, 20190556. https://doi.org/10.1098/rstb.2019.0556

Bernstein, B.E., Humphrey, E.L., Erlich, R.L., Schneider, R., Bouman, P., Liu, J.S., Kouzarides, T., Schreiber, S.L., 2002. Methylation of histone H3 Lys 4 in coding regions of active genes. Proceedings of the National Academy of Sciences of the United States of America. https://doi.org/10.1073/pnas.082249499

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S.L., Lander, E.S., 2006. A Bivalent Chromatin Structure Marks Key

Developmental Genes in Embryonic Stem Cells. Cell. https://doi.org/10.1016/j.cell.2006.02.041

Bessho, Y., Sakata, R., Komatsu, S., Shiota, K., Yamada, S., Kageyama, R., 2001. Dynamic expression and essential functions of Hes7 in somite segmentation. Genes Dev 15, 2642–2647. https://doi.org/10.1101/gad.930601

Bhanot, P., Brink, M., Samos, C.H., Hsieh, J.C., Wang, Y., Macke, J.P., Andrew, D., Nathans, J., Nusse, R., 1996. A new member of the frizzled family from Drosophila functions as a Wingless receptor. Nature 382, 225–230. https://doi.org/10.1038/382225a0

Bhattacharya, A., Deng, J.M., Zhang, Z., Behringer, R., de Crombrugghe, B., Maity, S.N., 2003. The B subunit of the CCAAT box binding transcription factor complex (CBF/NF-Y) is essential for early mouse development and cell proliferation. Cancer Res 63, 8167–8172.

Bi, W., Liu, Y., Guo, J., Lin, Z., Liu, J., Zhou, M., Wismeijer, D., Pathak, J.L., Wu, G., 2018. All-trans retinoic-acid inhibits heterodimeric bone morphogenetic protein 2/7-stimulated osteoclastogenesis, and resorption activity. Cell & Bioscience 8, 48. https://doi.org/10.1186/s13578-018-0246-y

Biechele, S., Cox, B.J., Rossant, J., 2011. Porcupine homolog is required for canonical Wnt signaling and gastrulation in mouse embryos. Dev Biol 355, 275–285. https://doi.org/10.1016/j.ydbio.2011.04.029

Bilic, J., Huang, Y.-L., Davidson, G., Zimmermann, T., Cruciat, C.-M., Bienz, M., Niehrs, C., 2007. Wnt induces LRP6 signalosomes and promotes dishevelled-dependent LRP6 phosphorylation. Science 316, 1619–1622. https://doi.org/10.1126/science.1137065

Binagui-Casas, A., Dias, A., Guillot, C., Metzis, V., Saunders, D., 2021. Building consensus in neuromesodermal research: Current advances and future biomedical perspectives. Curr Opin Cell Biol 73, 133–140. https://doi.org/10.1016/j.ceb.2021.08.003

Bisson, J.A., Mills, B., Paul Helt, J.-C., Zwaka, T.P., Cohen, E.D., 2015. Wnt5a and Wnt11 inhibit the canonical Wnt pathway and promote cardiac progenitor development via the Caspase-dependent degradation of AKT. Developmental Biology 398, 80–96. https://doi.org/10.1016/j.ydbio.2014.11.015

Boccia, L., Gamakharia, S., Coester, B., Whiting, L., Lutz, T.A., Le Foll, C., 2020. Amylin brain circuitry. Peptides 132, 170366. https://doi.org/10.1016/j.peptides.2020.170366

Bole-Feysot, C., Goffin, V., Edery, M., Binart, N., Kelly, P.A., 1998. Prolactin (PRL) and its receptor: actions, signal transduction pathways and phenotypes observed in PRL receptor knockout mice. Endocr Rev 19, 225–268. https://doi.org/10.1210/edrv.19.3.0334

Bothma, J.P., Garcia, H.G., Ng, S., Perry, M.W., Gregor, T., Levine, M., 2015. Enhancer additivity and non-additivity are determined by enhancer strength in the Drosophila embryo. eLife 4, e07956. https://doi.org/10.7554/eLife.07956

Boutros, M., Paricio, N., Strutt, D.I., Mlodzik, M., 1998. Dishevelled Activates JNK and Discriminates between JNK Pathways in Planar Polarity and wingless Signaling. Cell 94, 109–118. https://doi.org/10.1016/S0092-8674(00)81226-X

Boyle, A.P., Davis, S., Shulha, H.P., Meltzer, P., Margulies, E.H., Weng, Z., Furey, T.S., Crawford, G.E., 2008. High-resolution mapping and characterization of open chromatin across the genome. Cell 132, 311–322. https://doi.org/10.1016/j.cell.2007.12.014

Brand, N., Petkovich, M., Krust, A., Chambon, P., de Thé, H., Marchio, A., Tiollais, P., Dejean, A., 1988. Identification of a second human retinoic acid receptor. Nature 332, 850–853. https://doi.org/10.1038/332850a0

Brennan, J., Lu, C.C., Norris, D.P., Rodriguez, T.A., Beddington, R.S.P., Robertson, E.J., 2001. Nodal signalling in the epiblast patterns the early mouse embryo. Nature 411, 965–969. https://doi.org/10.1038/35082103

Brennan, J., Norris, D.P., Robertson, E.J., 2002. Nodal activity in the node governs left-right asymmetry. Genes Dev. 16, 2339–2344. https://doi.org/10.1101/gad.1016202

Brinkmann, E.-M., Mattes, B., Kumar, R., Hagemann, A.I.H., Gradl, D., Scholpp, S., Steinbeisser, H., Kaufmann, L.T., Özbek, S., 2016. Secreted Frizzled-related Protein 2 (sFRP2) Redirects Non-canonical Wnt Signaling from Fz7 to Ror2 during Vertebrate Gastrulation*. Journal of Biological Chemistry 291, 13730–13742. https://doi.org/10.1074/jbc.M116.733766

Brown, C.B., Engleka, K.A., Wenning, J., Min Lu, M., Epstein, J.A., 2005. Identification of a hypaxial somite enhancer element regulating Pax3 expression in migrating myoblasts and characterization of hypaxial muscle Cre transgenic mice. genesis 41, 202–209. https://doi.org/10.1002/gene.20116

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., Greenleaf, W.J., 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. https://doi.org/10.1038/nmeth.2688

Buenrostro, J.D., Wu, B., Litzenburger, U.M., Ruff, D., Gonzales, M.L., Snyder, M.P., Chang, H.Y., Greenleaf, W.J., 2015. Single-cell chromatin accessibility reveals principles of regulatory variation. Nature 523, 486–490. https://doi.org/10.1038/nature14590

Busada, J.T., Chappell, V.A., Niedenberger, B.A., Kaye, E.P., Keiper, B.D., Hogarth, C.A., Geyer, C.B., 2015. Retinoic acid regulates Kit translation during spermatogonial differentiation in the mouse. Developmental Biology 397, 140–149. https://doi.org/10.1016/j.ydbio.2014.10.020

Cambray, N., Wilson, V., 2007. Two distinct sources for a population of maturing axial progenitors. Development 134, 2829–2840. https://doi.org/10.1242/dev.02877

Cannavò, E., Khoueiry, P., Garfield, D.A., Geeleher, P., Zichner, T., Gustafson, E.H., Ciglar, L., Korbel, J.O., Furlong, E.E.M., 2016. Shadow Enhancers Are Pervasive Features of Developmental Regulatory Networks. Current Biology 26, 38–51. https://doi.org/10.1016/j.cub.2015.11.034

Casaca, A., Nóvoa, A., Mallo, M., 2016. Hoxb6 can interfere with somitogenesis in the posterior embryo through a mechanism independent of its rib-promoting activity. Development 143, 437–448. https://doi.org/10.1242/dev.133074

Chalamalasetty, R.B., Garriock, R.J., Dunty, W.C., Kennedy, M.W., Jailwala, P., Si, H., Yamaguchi, T.P., 2014. Mesogenin 1 is a master regulator of paraxial presomitic mesoderm differentiation. Development 141, 4285–4297. https://doi.org/10.1242/dev.110908

Chang, H.-K., Hou, W.-S., 2015. Retinoic Acid Modulates Interferon-γ Production by Hepatic Natural Killer T Cells via Phosphatase 2A and the Extracellular Signal-Regulated Kinase Pathway. Journal of Interferon & Cytokine Research 35, 200–212. https://doi.org/10.1089/jir.2014.0098

Chawengsaksophak, K., James, R., Hammond, V.E., Köntgen, F., Beck, F., 1997. Homeosis and intestinal tumours in Cdx2 mutant mice. Nature 386, 84–87. https://doi.org/10.1038/386084a0

Chen, J.D., Evans, R.M., 1995. A transcriptional co-repressor that interacts with nuclear hormone receptors. Nature 377, 454–457. https://doi.org/10.1038/377454a0

Chesley, P., 1935. Development of the short-tailed mutant in the house mouse. https://doi.org/10.1002/JEZ.1400700306

Chigurupati, S., Kulkarni, T., Thomas, S., Shah, G., 2005. Calcitonin stimulates multiple stages of angiogenesis by directly acting on endothelial cells. Cancer Res 65, 8519–8529. https://doi.org/10.1158/0008-5472.CAN-05-0848

Choi, J., Park, S.Y., Costantini, F., Jho, E., Joo, C.-K., 2004. Adenomatous Polyposis Coli Is Down-regulated by the Ubiquitin-Proteasome Pathway in a Process Facilitated by Axin*. Journal of Biological Chemistry 279, 49188–49198. https://doi.org/10.1074/jbc.M404655200

Ciruna, B., Rossant, J., 2001. FGF Signaling Regulates Mesoderm Cell Fate Specification and Morphogenetic Movement at the Primitive Streak. Developmental Cell 1, 37–49. https://doi.org/10.1016/S1534-5807(01)00017-X

Clark, D.J., Felsenfeld, G., 1992. A nucleosome core is transferred out of the path of a transcribing polymerase. Cell 71, 11–22. https://doi.org/10.1016/0092-8674(92)90262-b

Cohen, E.D., Miller, M.F., Wang, Z., Moon, R.T., Morrisey, E.E., 2012. Wnt5a and Wnt11 are essential for second heart field progenitor development. Development 139, 1931–1940. https://doi.org/10.1242/dev.069377

Cong, F., Schweizer, L., Varmus, H., 2004. Wnt signals across the plasma membrane to activate the beta-catenin pathway by forming oligomers containing its receptors, Frizzled and LRP. Development 131, 5103–5115. https://doi.org/10.1242/dev.01318

Conlon, F.L., Barth, K.S., Robertson, E.J., 1991. A novel retrovirally induced embryonic lethal mutation in the mouse: assessment of the developmental fate of embryonic stem cells homozygous for the 413.d proviral integration. Development 111, 969–981. https://doi.org/10.1242/dev.111.4.969

Coombs, G.S., Yu, J., Canning, C.A., Veltri, C.A., Covey, T.M., Cheong, J.K., Utomo, V., Banerjee, N., Zhang, Z.H., Jadulco, R.C., Concepcion, G.P., Bugni, T.S., Harper, M.K., Mihalek, I., Jones, C.M., Ireland, C.M., Virshup, D.M., 2010. WLS-dependent secretion of WNT3A requires Ser209 acylation and vacuolar acidification. Journal of Cell Science 123, 3357–3367. https://doi.org/10.1242/jcs.072132

Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., Kathiria, A., Cho, S.W., Mumbach, M.R., Carter, A.C., Kasowski, M., Orloff, L.A., Risca, V.I., Kundaje, A., Khavari, P.A., Montine, T.J., Greenleaf, W.J., Chang, H.Y., 2017. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat Methods 14, 959–962. https://doi.org/10.1038/nmeth.4396

Correia, Rion B., Almeida, J.M., Wyrwoll, M.J., Julca, I., Sobral, D., Misra, C.S., Guilgur, L.G., Schuppe, H.-C., Silva, N., Prudêncio, P., Nóvoa, A., Leocádio, A.S., Bom, J., Mallo, M., Kliesch, S., Mutwil, M., Rocha, L.M., Tüttelmann, F., Becker, J.D., Navarro-Costa, P., 2022. The conserved transcriptional program of metazoan male germ cells uncovers ancient origins of human infertility. https://doi.org/10.1101/2022.03.02.482557

Correia, Rion B., Barrat, A., Rocha, L.M., 2022. The metric backbone preserves community structure and is a primary transmission subgraph in contact networks. https://doi.org/10.1101/2022.02.02.478784

Correia, R.B., Costa, P.N., Rocha, L.M., 2020. Extraction of overlapping modules in networks via spectral methods and information theory. Presented at the The 9th International Workshop on Complex Networks and Their Applications, Madrid, Spain.

Coucouvanis, E., Martin, G.R., 1995. Signals for death and survival: A two-step mechanism for cavitation in the vertebrate embryo. Cell 83, 279–287. https://doi.org/10.1016/0092-8674(95)90169-8

Crawford, G.E., Davis, S., Scacheri, P.C., Renaud, G., Halawi, M.J., Erdos, M.R., Green, R., Meltzer, P.S., Wolfsberg, T.G., Collins, F.S., 2006. DNase-chip: a high-resolution method to identify DNase I hypersensitive sites using tiled microarrays. Nat Methods 3, 503–509. https://doi.org/10.1038/nmeth888

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., Boyer, L.A., Young, R.A., Jaenisch, R., 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences of the United States of America. https://doi.org/10.1073/pnas.1016071107

Cunningham, T.J., Brade, T., Sandell, L.L., Lewandoski, M., Trainor, P.A., Colas, A., Mercola, M., Duester, G., 2015a. Retinoic Acid Activity in Undifferentiated Neural Progenitors Is Sufficient to Fulfill Its Role in Restricting Fgf8 Expression for Somitogenesis. PLOS ONE 10, e0137894. https://doi.org/10.1371/journal.pone.0137894

Cunningham, T.J., Colas, A., Duester, G., 2016. Early molecular events during retinoic acid induced differentiation of neuromesodermal progenitors. Biology Open 5, 1821–1833. https://doi.org/10.1242/bio.020891

Cunningham, T.J., Kumar, S., Yamaguchi, T.P., Duester, G., 2015b. Wnt8a and Wnt3a Cooperate in the Axial Stem Cell Niche to Promote Mammalian Body Axis Extension. Dev Dyn 244, 797–807. https://doi.org/10.1002/dvdy.24275

Cunningham, T.J., Lancman, J.J., Berenguer, M., Dong, P.D.S., Duester, G., 2018. Genomic Knockout of Two Presumed Forelimb Tbx5 Enhancers Reveals They Are Nonessential for Limb Development. Cell Rep 23, 3146–3151. https://doi.org/10.1016/j.celrep.2018.05.052

Cusanovich, D.A., Daza, R., Adey, A., Pliner, H.A., Christiansen, L., Gunderson, K.L., Steemers, F.J., Trapnell, C., Shendure, J., 2015. Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. Science 348, 910–914. https://doi.org/10.1126/science.aab1601

Dawson, M.I., Xia, Z., 2012. The Retinoid X Receptors and Their Ligands. Biochim Biophys Acta 1821, 21–56. https://doi.org/10.1016/j.bbalip.2011.09.014

Deal, R.B., Henikoff, J.G., Henikoff, S., 2010. Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. Science 328, 1161–1164. https://doi.org/10.1126/science.1186777

Deema, 2019. Image retrieved from: https://commons.wikimedia.org/wiki/File:ATAC-Seq_Figure_.svg.

DeVeale, B., Brokhman, I., Mohseni, P., Babak, T., Yoon, C., Lin, A., Onishi, K., Tomilin, A., Pevny, L., Zandstra, P.W., Nagy, A., Kooy, D. van der, 2013. Oct4 Is Required ~E7.5 for Proliferation in the Primitive Streak. PLOS Genetics 9, e1003957. https://doi.org/10.1371/journal.pgen.1003957

Dey, N., De, P.K., Wang, M., Zhang, H., Dobrota, E.A., Robertson, K.A., Durden, D.L., 2007. CSK Controls Retinoic Acid Receptor (RAR) Signaling: a RAR-c-SRC Signaling Axis Is Required for Neuritogenic Differentiation. Mol Cell Biol 27, 4179–4197. https://doi.org/10.1128/MCB.01352-06

Dias, A., Lozovska, A., Wymeersch, F.J., Nóvoa, A., Binagui-Casas, A., Sobral, D., Martins, G.G., Wilson, V., Mallo, M., 2020. A Tgfbr1/Snai1-dependent developmental module at the core of vertebrate axial elongation. eLife 9, e56615. https://doi.org/10.7554/eLife.56615

Diaz-Cuadros, M., Miettinen, T.P., Skinner, O.S., Sheedy, D., Díaz-García, C.M., Gapon, S., Hubaud, A., Yellen, G., Manalis, S.R., Oldham, W.M., Pourquié, O., 2023. Metabolic regulation of species-specific developmental rates. Nature 613, 550–557. https://doi.org/10.1038/s41586-022-05574-4

Diaz-Cuadros, M., Pourquie, O., 2021. In vitro systems: A new window to the segmentation clock. Dev Growth Differ 63, 140–153. https://doi.org/10.1111/dgd.12710

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. https://doi.org/10.1016/s0896-6273(03)00565-8

Ding, J., Yang, L., Yan, Y.T., Chen, A., Desai, N., Wynshaw-Boris, A., Shen, M.M., 1998. Cripto is required for correct orientation of the anterior-posterior axis in the mouse embryo. Nature 395, 702–707. https://doi.org/10.1038/27215

Divate, M., Cheung, E., 2018. GUAVA: A Graphical User Interface for the Analysis and Visualization of ATAC-seq Data. Frontiers in Genetics 9.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., Ren, B., 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485, 376–380. https://doi.org/10.1038/nature11082

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T.R., 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. https://doi.org/10.1093/bioinformatics/bts635

Dong, D., Ruuska, S.E., Levinthal, D.J., Noy, N., 1999. Distinct Roles for Cellular Retinoic Acid-binding Proteins I and II in Regulating Signaling by Retinoic Acid *. Journal of Biological Chemistry 274, 23695–23698. https://doi.org/10.1074/jbc.274.34.23695

Dorrello, N.V., Peschiaroli, A., Guardavaccaro, D., Colburn, N.H., Sherman, N.E., Pagano, M., 2006. S6K1- and ßTRCP-Mediated Degradation of PDCD4 Promotes Protein Translation and Cell Growth. Science 314, 467–471. https://doi.org/10.1126/science.1130276

Doubravska, L., Krausova, M., Gradl, D., Vojtechova, M., Tumova, L., Lukas, J., Valenta, T., Pospichalova, V., Fafilek, B., Plachy, J., Sebesta, O., Korinek, V., 2011. Fatty acid modification of Wnt1 and Wnt3a at serine is prerequisite for lipidation at cysteine and is essential for Wnt signalling. Cellular Signalling 23, 837–848. https://doi.org/10.1016/j.cellsig.2011.01.007

Downs, K.M., Davies, T., 1993. Staging of gastrulating mouse embryos by morphological landmarks in the dissecting microscope. Development 118, 1255–1266. https://doi.org/10.1242/dev.118.4.1255

Duboule, D., Dollé, P., 1989. The structural and functional organization of the murine HOX gene family resembles that of Drosophila homeotic genes. EMBO J 8, 1497–1505. https://doi.org/10.1002/j.1460-2075.1989.tb03534.x

Dubrulle, J., McGrew, MJ., Pourquié, O., 2001. FGF signaling controls somite boundary position and regulates segmentation clock control of spatiotemporal Hox gene activation. Cell 106. https://doi.org/10.1016/s0092-8674(01)00437-8

Dupé, V., Matt, N., Garnier, J.-M., Chambon, P., Mark, M., Ghyselinck, N.B., 2003. A newborn lethal defect due to inactivation of retinaldehyde dehydrogenase type 3 is prevented by maternal retinoic acid treatment. Proc Natl Acad Sci U S A 100, 14036–14041. https://doi.org/10.1073/pnas.2336223100

Economides, K.D., Zeltser, L., Capecchi, M.R., 2003. Hoxb13 mutations cause overgrowth of caudal spinal cordand tail vertebrae. Developmental Biology 256, 317–330. https://doi.org/10.1016/S0012-1606(02)00137-9

El Haddad, N., Moore, R., Heathcote, D., Mounayar, M., Azzi, J., Mfarrej, B., Batal, I., Ting, C., Atkinson, M., Sayegh, M.H., Ashton-Rickardt, P.G., Abdi, R., 2011.

The novel role of SERPINB9 in cytotoxic protection of human mesenchymal stem cells. J Immunol 187, 2252–2260. https://doi.org/10.4049/jimmunol.1003981

El-Sherif, E., Levine, M., 2016. Shadow Enhancers Mediate Dynamic Shifts of Gap Gene Expression in the Drosophila Embryo. Current Biology 26, 1164–1169. https://doi.org/10.1016/j.cub.2016.02.054

England, S., Batista, M.F., Mich, J.K., Chen, J.K., Lewis, K.E., 2011. Roles of Hedgehog pathway components and retinoic acid signalling in specifying zebrafish ventral spinal cord neurons. Development 138, 5121–5134. https://doi.org/10.1242/dev.066159

Fan, X., Molotkov, A., Manabe, S.-I., Donmoyer, C.M., Deltour, L., Foglio, M.H., Cuenca, A.E., Blaner, W.S., Lipton, S.A., Duester, G., 2003. Targeted Disruption of Aldh1a1 (Raldh1) Provides Evidence for a Complex Mechanism of Retinoic Acid Synthesis in the Developing Retina. Mol Cell Biol 23, 4637–4648. https://doi.org/10.1128/MCB.23.13.4637-4648.2003

Fanto, M., Weber, U., Strutt, D.I., Mlodzik, M., 2000. Nuclear signaling by Rac and Rho GTPases is required in the establishment of epithelial planar polarity in the Drosophila eye. Current Biology 10, S1. https://doi.org/10.1016/S0960-9822(00)00645-X

Farese, R.V., Ruland, S.L., Flynn, L.M., Stokowski, R.P., Young, S.G., 1995. Knockout of the mouse apolipoprotein B gene results in embryonic lethality in homozygotes and protection against diet-induced hypercholesterolemia in heterozygotes. Proceedings of the National Academy of Sciences 92, 1774–1778. https://doi.org/10.1073/pnas.92.5.1774

Ferretti, E., Hadjantonakis, A.-K., 2019. Mesoderm specification and diversification: from single cells to emergent tissues. Curr Opin Cell Biol 61, 110–116. https://doi.org/10.1016/j.ceb.2019.07.012

Frankel, N., Davis, G.K., Vargas, D., Wang, S., Payre, F., Stern, D.L., 2010. Phenotypic robustness conferred by apparently redundant transcriptional enhancers. Nature 466, 490–493. https://doi.org/10.1038/nature09158

Fromental-Ramain, C., Warot, X., Lakkaraju, S., Favier, B., Haack, H., Birling, C., Dierich, A., Doll e, P., Chambon, P., 1996. Specific and redundant functions of the paralogous Hoxa-9 and Hoxd-9 genes in forelimb and axial skeleton patterning. Development 122, 461–472. https://doi.org/10.1242/dev.122.2.461

Gabriele, M., Brandão, H.B., Grosse-Holz, S., Jha, A., Dailey, G.M., Cattoglio, C., Hsieh, T.-H.S., Mirny, L., Zechner, C., Hansen, A.S., 2022. Dynamics of CTCF- and cohesin-mediated chromatin looping revealed by live-cell imaging. Science 376, 496–501. https://doi.org/10.1126/science.abn6583

Gallet, A., Rodriguez, R., Ruel, L., Therond, P.P., 2003. Cholesterol modification of hedgehog is required for trafficking and movement, revealing an asymmetric cellular response to hedgehog. Dev Cell 4, 191–204. https://doi.org/10.1016/s1534-5807(03)00031-5

Galli, L.M., Burrus, L.W., 2011. Differential Palmit(e)oylation of Wnt1 on C93 and S224 Residues Has Overlapping and Distinct Consequences. PLOS ONE 6, e26636. https://doi.org/10.1371/journal.pone.0026636

Gavalas, A., Ruhrberg, C., Livet, J., Henderson, C.E., Krumlauf, R., 2003. Neuronal defects in the hindbrain of Hoxa1, Hoxb1 and Hoxb2 mutants reflect regulatory interactions among these Hox genes. Development 130, 5663–5679. https://doi.org/10.1242/dev.00802

Gentile, C., Berlivet, S., Mayran, A., Paquette, D., Guerard-Millet, F., Bajon, E., Dostie, J., Kmita, M., 2019. PRC2-Associated Chromatin Contacts in the Developing Limb Reveal a Possible Mechanism for the Atypical Role of PRC2 in HoxA Gene Expression. Dev Cell 50, 184-196.e4. https://doi.org/10.1016/j.devcel.2019.05.021

Giguere, V., Ong, E.S., Segui, P., Evans, R.M., 1987. Identification of a receptor for the morphogen retinoic acid. Nature 330, 624–629. https://doi.org/10.1038/330624a0

Giresi, P.G., Kim, J., McDaniell, R.M., Iyer, V.R., Lieb, J.D., 2007. FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. Genome Res 17, 877–885. https://doi.org/10.1101/gr.5533506

Golden-Mason, L., McMahan, R.H., Strong, M., Reisdorph, R., Mahaffey, S., Palmer, B.E., Cheng, L., Kulesza, C., Hirashima, M., Niki, T., Rosen, H.R., 2013. Galectin-9 functionally impairs natural killer cells in humans and mice. J Virol 87, 4835–4845. https://doi.org/10.1128/JVI.01085-12

Gomez, C., Özbudak, E.M., Wunderlich, J., Baumann, D., Lewis, J., Pourquié, O., 2008. Control of segment number in vertebrate embryos. Nature 454, 335–339. https://doi.org/10.1038/nature07020

Goulding, M.D., Chalepakis, G., Deutsch, U., Erselius, J.R., Gruss, P., 1991. Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. EMBO J 10, 1135–1147. https://doi.org/10.1002/j.1460-2075.1991.tb08054.x

Gouti, M., Delile, J., Stamataki, D., Wymeersch, F.J., Huang, Y., Kleinjung, J., Wilson, V., Briscoe, J., 2017. A Gene Regulatory Network Balances Neural and Mesoderm Specification during Vertebrate Trunk Development. Developmental Cell 41, 243-261.e7. https://doi.org/10.1016/j.devcel.2017.04.002

Graham, A., Papalopulu, N., Krumlauf, R., 1989. The murine and Drosophila homeobox gene complexes have common features of organization and expression. Cell 57, 367–378. https://doi.org/10.1016/0092-8674(89)90912-4

Graham, V., Khudyakov, J., Ellis, P., Pevny, L., 2003. SOX2 Functions to Maintain Neural Progenitor Identity. Neuron 39, 749–765. https://doi.org/10.1016/S0896-6273(03)00497-5

Gupta, P., Ho, P.-C., Huq, M.M., Ha, S.G., Park, S.W., Khan, A.A., Tsai, N.-P., Wei, L.-N., 2008. Retinoic acid-stimulated sequential phosphorylation, PML recruitment, and SUMOylation of nuclear receptor TR2 to suppress Oct4

expression. Proceedings of the National Academy of Sciences 105, 11424–11429. https://doi.org/10.1073/pnas.0710561105

Han, H., Cortez, C.C., Yang, X., Nichols, P.W., Jones, P.A., Liang, G., 2011. DNA methylation directly silences genes with non-CpG island promoters and establishes a nucleosome occupied promoter. Hum Mol Genet 20, 4299–4310. https://doi.org/10.1093/hmg/ddr356

Hart, M., Concordet, J.-P., Lassot, I., Albert, I., del los Santos, R., Durand, H., Perret, C., Rubinfeld, B., Margottin, F., Benarous, R., Polakis, P., 1999. The F-box protein β-TrCP associates with phosphorylated β-catenin and regulates its activity in the cell. Current Biology 9, 207–211. https://doi.org/10.1016/S0960-9822(99)80091-8

Haussler, M.R., Whitfield, G.K., Kaneko, I., Haussler, C.A., Hsieh, D., Hsieh, J.-C., Jurutka, P.W., 2013. Molecular Mechanisms of Vitamin D Action. Calcif Tissue Int 92, 77–98. https://doi.org/10.1007/s00223-012-9619-0

Hebbes, T.R., Clayton, A.L., Thorne, A.W., Crane-Robinson, C., 1994. Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain. EMBO J 13, 1823–1830. https://doi.org/10.1002/j.1460-2075.1994.tb06451.x

Heitz, E., 1928. Das Heterochromatin der Moose. Jahrb Wiss Bot. 762–818.

Hempel, K., Lange, H.W., Birkofer, L., 1968. [Epsilon-N-trimethyllysine, a new amino acid in histones]. Naturwissenschaften 55, 37. https://doi.org/10.1007/BF00593411

Hernández, A.R., Klein, A.M., Kirschner, M.W., 2012. Kinetic responses of β-catenin specify the sites of Wnt control. Science 338, 1337–1340. https://doi.org/10.1126/science.1228734

Hernandez, R.E., Putzke, A.P., Myers, J.P., Margaretha, L., Moens, C.B., 2007. Cyp26 enzymes generate the retinoic acid response pattern necessary for hindbrain development. Development 134, 177–187. https://doi.org/10.1242/dev.02706

Herr, P., Basler, K., 2012. Porcupine-mediated lipidation is required for Wnt recognition by Wls. Developmental Biology 361, 392–402. https://doi.org/10.1016/j.ydbio.2011.11.003

Herrmann, B.G., Labeit, S., Poustka, A., King, T.R., Lehrach, H., 1990. Cloning of the T gene required in mesoderm formation in the mouse. Nature 343, 617–622. https://doi.org/10.1038/343617a0

Ho, H.-Y.H., Susman, M.W., Bikoff, J.B., Ryu, Y.K., Jonas, A.M., Hu, L., Kuruvilla, R., Greenberg, M.E., 2012. Wnt5a–Ror–Dishevelled signaling constitutes a core developmental pathway that controls tissue morphogenesis. Proceedings of the National Academy of Sciences 109, 4044–4051. https://doi.org/10.1073/pnas.1200421109

Hogan, B., Beddington, R., Constantini, F., 1994. Manipulating the Mouse Embryo: A Laboratory Manual, 4th ed. Cold Spring Harbor.

Hong, S.-H., Fang, S., Lu, B.C., Nofsinger, R., Kawakami, Y., Castro, G.L., Yin, Y., Lin, C., Yu, R.T., Downes, M., Izpisúa Belmonte, J.C., Shilatifard, A., Evans, R.M., 2018. Corepressor SMRT is required to maintain Hox transcriptional memory

during somitogenesis. Proceedings of the National Academy of Sciences 115, 10381–10386. https://doi.org/10.1073/pnas.1809480115

Horan, G.S., Ramírez-Solis, R., Featherstone, M.S., Wolgemuth, D.J., Bradley, A., Behringer, R.R., 1995. Compound mutants for the paralogous hoxa-4, hoxb-4, and hoxd-4 genes show more complete homeotic transformations and a dose-dependent increase in the number of vertebrae transformed. Genes Dev. 9, 1667–1677. https://doi.org/10.1101/gad.9.13.1667

Houle, M., Prinos, P., Iulianella, A., Bouchard, N., Lohnes, D., 2000. Retinoic acid regulation of Cdx1: an indirect mechanism for retinoids and vertebral specification. Mol Cell Biol 20, 6579–6586. https://doi.org/10.1128/MCB.20.17.6579-6586.2000

Houle, M., Sylvestre, J.-R., Lohnes, D., 2003. Retinoic acid regulates a subset of Cdx1 function in vivo. Development 130, 6555–6567. https://doi.org/10.1242/dev.00889

Ikeda, S., Kishida, S., Yamamoto, H., Murai, H., Koyama, S., Kikuchi, A., 1998. Axin, a negative regulator of the Wnt signaling pathway, forms a complex with GSK-3beta and beta-catenin and promotes GSK-3beta-dependent phosphorylation of beta-catenin. EMBO J 17, 1371–1384. https://doi.org/10.1093/emboj/17.5.1371

Ikeya, M., Lee, S.M.K., Johnson, J.E., McMahon, A.P., Takada, S., 1997. Wnt signalling required for expansion of neural crest and CNS progenitors. Nature 389, 966–970. https://doi.org/10.1038/40146

Irie, N., Satoh, N., Kuratani, S., 2018. The phylum Vertebrata: a case for zoological recognition. Zoological Letters 4, 32. https://doi.org/10.1186/s40851-018-0114-y

Iwafuchi-Doi, M., Donahue, G., Kakumanu, A., Watts, J.A., Mahony, S., Pugh, B.F., Lee, D., Kaestner, K.H., Zaret, K.S., 2016. The pioneer transcription factor FoxA maintains an accessible nucleosome configuration at enhancers for tissue-specific gene activation. Mol Cell 62, 79–91. https://doi.org/10.1016/j.molcel.2016.03.001

Izpisúa-Belmonte, J.C., Falkenstein, H., Dollé, 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. https://doi.org/10.1002/j.1460-2075.1991.tb07764.x

Jaffe, L.F., 1981. The role of ionic currents in establishing developmental pattern. Philos Trans R Soc Lond B Biol Sci 295, 553–566. https://doi.org/10.1098/rstb.1981.0160

Jerković, I., Ibrahim, D.M., Andrey, G., Haas, S., Hansen, P., Janetzki, C., González Navarrete, I., Robinson, P.N., Hecht, J., Mundlos, S., 2017. Genome-Wide Binding of Posterior HOXA/D Transcription Factors Reveals Subgrouping and Association with CTCF. PLoS Genet 13, e1006567. https://doi.org/10.1371/journal.pgen.1006567

Jolma, A., Yan, J., Whitington, T., Toivonen, J., Nitta, K.R., Rastas, P., Morgunova, E., Enge, M., Taipale, M., Wei, G., Palin, K., Vaquerizas, J.M., Vincentelli, R.,

Luscombe, N.M., Hughes, T.R., Lemaire, P., Ukkonen, E., Kivioja, T., Taipale, J., 2013. DNA-binding specificities of human transcription factors. Cell 152, 327–339. https://doi.org/10.1016/j.cell.2012.12.009

Jurberg, A.D., Aires, R., Nóvoa, A., Rowland, J.E., Mallo, M., 2014. Compartment-dependent activities of Wnt3a/β-catenin signaling during vertebrate axial extension. Dev Biol 394, 253–263. https://doi.org/10.1016/j.ydbio.2014.08.012

Jurberg, A.D., Aires, R., Varela-Lasheras, I., Nóvoa, A., Mallo, M., 2013. Switching Axial Progenitors from Producing Trunk to Tail Tissues in Vertebrate Embryos. Developmental Cell 25, 451–462. https://doi.org/10.1016/j.devcel.2013.05.009

Kaiser, K., Gyllborg, D., Procházka, J., Salašová, A., Kompaníková, P., Molina, F.L., Laguna-Goya, R., Radaszkiewicz, T., Harnoš, J., Procházková, M., Potěšil, D., Barker, R.A., Casado, Á.G., Zdráhal, Z., Sedláček, R., Arenas, E., Villaescusa, J.C., Bryja, V., 2019. WNT5A is transported via lipoprotein particles in the cerebrospinal fluid to regulate hindbrain morphogenesis. Nat Commun 10, 1498. https://doi.org/10.1038/s41467-019-09298-4

Kaname, T., Kuwano, A., Murano, I., Uehara, K., Muramatsu, T., Kajii, T., 1993. Midkine gene (MDK), a gene for prenatal differentiation and neuroregulation, maps to band 11p11.2 by fluorescence in situ hybridization. Genomics 17, 514–515. https://doi.org/10.1006/geno.1993.1359

Kemp, C., Willems, E., Abdo, S., Lambiv, L., Leyns, L., 2005. Expression of all Wnt genes and their secreted antagonists during mouse blastocyst and postimplantation development. Developmental Dynamics 233, 1064–1075. https://doi.org/10.1002/dvdy.20408

Kennedy, K.A.M., Porter, T., Mehta, V., Ryan, S.D., Price, F., Peshdary, V., Karamboulas, C., Savage, J., Drysdale, T.A., Li, S.-C., Bennett, S.A.L., Skerjanc, I.S., 2009. Retinoic acid enhances skeletal muscle progenitor formation and bypasses inhibition by bone morphogenetic protein 4 but not dominant negative beta-catenin. BMC Biol 7, 67. https://doi.org/10.1186/1741-7007-7-67

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., Haussler, and D., 2002. The Human Genome Browser at UCSC. Genome Res. 12, 996–1006. https://doi.org/10.1101/gr.229102

Kiani, K., Sanford, E.M., Goyal, Y., Raj, A., 2022. Changes in chromatin accessibility are not concordant with transcriptional changes for single-factor perturbations. Mol Syst Biol 18, e10979. https://doi.org/10.15252/msb.202210979

Kieffer-Kwon, K.-R., Tang, Z., Mathe, E., Qian, J., Sung, M.-H., Li, G., Resch, W., Baek, S., Pruett, N., Grøntved, L., Vian, L., Nelson, S., Zare, H., Hakim, O., Reyon, D., Yamane, A., Nakahashi, H., Kovalchuk, A.L., Zou, J., Joung, J.K., Sartorelli, V., Wei, C.-L., Ruan, X., Hager, G.L., Ruan, Y., Casellas, R., 2013. Interactome Maps of Mouse Gene Regulatory Domains Reveal Basic Principles of Transcriptional Regulation. Cell 155, 1507–1520. https://doi.org/10.1016/j.cell.2013.11.039

Kim, S., Jho, E., 2010. The Protein Stability of Axin, a Negative Regulator of Wnt Signaling, Is Regulated by Smad Ubiquitination Regulatory Factor 2 (Smurf2)*.

Journal of Biological Chemistry 285, 36420–36426. https://doi.org/10.1074/jbc.M110.137471

Kinder, S.J., Tsang, T.E., Quinlan, G.A., Hadjantonakis, A.K., Nagy, A., Tam, P.P., 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. https://doi.org/10.1242/dev.126.21.4691

Kirn-Safran, C.B., Oristian, D.S., Focht, R.J., Parker, S.G., Vivian, J.L., Carson, D.D., 2007. Global growth deficiencies in mice lacking the ribosomal protein HIP/RPL29. Dev Dyn 236, 447–460. https://doi.org/10.1002/dvdy.21046

Kitagawa, M., Hatakeyama, S., 1999. An F-box protein, FWD1, mediates ubiquitin-dependent proteolysis of β-catenin. The EMBO Journal 18, 2401–2410. https://doi.org/10.1093/emboj/18.9.2401

Kitagawa, M., Hatakeyama, S., Shirane, M., Matsumoto, M., Ishida, N., Hattori, K., Nakamichi, I., Kikuchi, A., Nakayama, K., Nakayama, K., 1999. An F-box protein, FWD1, mediates ubiquitin-dependent proteolysis of beta-catenin. EMBO J 18, 2401–2410. https://doi.org/10.1093/emboj/18.9.2401

Koch, F., Scholze, M., Wittler, L., Schifferl, D., Sudheer, S., Grote, P., Timmermann, B., Macura, K., Herrmann, B.G., 2017. Antagonistic Activities of Sox2 and Brachyury Control the Fate Choice of Neuro-Mesodermal Progenitors. Developmental Cell 42, 514-526.e7. https://doi.org/10.1016/j.devcel.2017.07.021

Koo, B.-K., Spit, M., Jordens, I., Low, T.Y., Stange, D.E., van de Wetering, M., van Es, J.H., Mohammed, S., Heck, A.J.R., Maurice, M.M., Clevers, H., 2012. Tumour suppressor RNF43 is a stem-cell E3 ligase that induces endocytosis of Wnt receptors. Nature 488, 665–669. https://doi.org/10.1038/nature11308

Krawchuk, D., Kania, A., 2008. Identification of genes controlled by LMX1B in the developing mouse limb bud. Developmental Dynamics 237, 1183–1192. https://doi.org/10.1002/dvdy.21514

Krol, A.J., Roellig, D., Dequéant, M.-L., Tassy, O., Glynn, E., Hattem, G., Mushegian, A., Oates, A.C., Pourquié, O., 2011. Evolutionary plasticity of segmentation clock networks. Development 138, 2783–2792. https://doi.org/10.1242/dev.063834

Kühl, M., Sheldahl, L.C., Malbon, C.C., Moon, R.T., 2000. Ca2+/Calmodulin-dependent Protein Kinase II Is Stimulated by Wnt and Frizzled Homologs and Promotes Ventral Cell Fates in Xenopus*. Journal of Biological Chemistry 275, 12701–12711. https://doi.org/10.1074/jbc.275.17.12701

Kulakovskiy, I.V., Vorontsov, I.E., Yevshin, I.S., Sharipov, R.N., Fedorova, A.D., Rumynskiy, E.I., Medvedeva, Y.A., Magana-Mora, A., Bajic, V.B., Papatsenko, D.A., Kolpakov, F.A., Makeev, V.J., 2018. HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. Nucleic Acids Research 46, D252–D259. https://doi.org/10.1093/nar/gkx1106

Kumar, A., Lualdi, M., Lewandoski, M., Kuehn, M.R., 2008. Broad mesodermal and endodermal deletion of Nodal at postgastrulation stages results solely in left/right axial defects. Developmental Dynamics 237, 3591–3601. https://doi.org/10.1002/dvdy.21665

Kumar, S., Cunningham, T.J., Duester, G., 2016. Nuclear receptor corepressors Ncor1 and Ncor2 (Smrt) are required for retinoic acid-dependent repression of Fgf8 during somitogenesis. Dev Biol 418, 204–215. https://doi.org/10.1016/j.ydbio.2016.08.005

Kumar, S., Duester, G., 2014. Retinoic acid controls body axis extension by directly repressing Fgf8 transcription. Development 141, 2972–2977. https://doi.org/10.1242/dev.112367

Kurokawa, R., Söderström, M., Hörlein, A., Halachmi, S., Brown, M., Rosenfeld, M.G., Glass, C.K., 1995. Polarity-specific activities of retinoic acid receptors determined by a co-repressor. Nature 377, 451–454. https://doi.org/10.1038/377451a0

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lönnerberg, P., Furlan, A., Fan, J., Borm, L.E., Liu, Z., van Bruggen, D., Guo, J., He, X., Barker, R., Sundström, E., Castelo-Branco, G., Cramer, P., Adameyko, I., Linnarsson, S., Kharchenko, P.V., 2018. RNA velocity of single cells. Nature 560, 494–498. https://doi.org/10.1038/s41586-018-0414-6

Lacomme, M., Medevielle, F., Bourbon, H.-M., Thierion, E., Kleinjan, D.-J., Roussat, M., Pituello, F., Bel-Vialar, S., 2018. A long range distal enhancer controls temporal fine-tuning of PAX6 expression in neuronal precursors. Developmental Biology 436, 94–107. https://doi.org/10.1016/j.ydbio.2018.02.015

Lanigan, T.M., Tverberg, L.A., Russo, A.F., 1993. Retinoic acid repression of cell-specific helix-loop-helix-octamer activation of the calcitonin/calcitonin gene-related peptide enhancer. Molecular and Cellular Biology 13, 6079–6088. https://doi.org/10.1128/mcb.13.10.6079-6088.1993

Lawson, K.A., Meneses, J.J., Pedersen, R.A., 1991. Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. Development 113, 891–911. https://doi.org/10.1242/dev.113.3.891

Lebensohn, A.M., Rohatgi, R., 2018. R-spondins can potentiate WNT signaling without LGRs. eLife 7, e33126. https://doi.org/10.7554/eLife.33126

Lee, H.-Y., Walsh, G.L., Dawson, M.I., Hong, W.K., Kurie, J.M., 1998. All-trans-Retinoic Acid Inhibits Jun N-terminal Kinase-dependent Signaling Pathways *. Journal of Biological Chemistry 273, 7066–7071. https://doi.org/10.1074/jbc.273.12.7066

Lee, K., Hsiung, C.C.-S., Huang, P., Raj, A., Blobel, G.A., 2015. Dynamic enhancer–gene body contacts during transcription elongation. Genes Dev. 29, 1992–1997. https://doi.org/10.1101/gad.255265.114

Lettice, L.A., Heaney, S.J.H., Purdie, L.A., Li, L., de Beer, P., Oostra, B.A., Goode, D., Elgar, G., Hill, R.E., de Graaff, E., 2003. A long-range Shh enhancer regulates

expression in the developing limb and fin and is associated with preaxial polydactyly. Hum Mol Genet 12, 1725–1735. https://doi.org/10.1093/hmg/ddg180

Levin, M., 2021. Bioelectric signaling: Reprogrammable circuits underlying embryogenesis, regeneration, and cancer. Cell 184, 1971–1989. https://doi.org/10.1016/j.cell.2021.02.034

Levin, M., Thorlin, T., Robinson, K.R., Nogi, T., Mercola, M., 2002. Asymmetries in H+/K+-ATPase and cell membrane potentials comprise a very early step in left-right patterning. Cell 111, 77–89. https://doi.org/10.1016/s0092-8674(02)00939-x

Li, A., He, M., Wang, H., Qiao, B., Chen, P., Gu, H., Zhang, M., He, S., 2007. All-trans retinoic acid negatively regulates cytotoxic activities of nature killer cell line 92. Biochemical and Biophysical Research Communications 352, 42–47. https://doi.org/10.1016/j.bbrc.2006.10.132

Li, Y., He, Y., Liang, Z., Wang, Y., Chen, F., Djekidel, M.N., Li, G., Zhang, X., Xiang, S., Wang, Z., Gao, J., Zhang, M.Q., Chen, Y., 2018. Alterations of specific chromatin conformation affect ATRA-induced leukemia cell differentiation. Cell Death Dis 9, 1–15. https://doi.org/10.1038/s41419-017-0173-6

Li, Z., Schulz, M.H., Look, T., Begemann, M., Zenke, M., Costa, I.G., 2019. Identification of transcription factor binding sites using ATAC-seq. Genome Biology 20, 45. https://doi.org/10.1186/s13059-019-1642-2

Lickert, H., Kispert, A., Kutsch, S., Kemler, R., 2001. Expression patterns of Wnt genes in mouse gut development. Mechanisms of Development, Molecular Mechanisms of Cell Migration during Embryogenesis 105, 181–184. https://doi.org/10.1016/S0925-4773(01)00390-2

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., Sandstrom, R., Bernstein, B., Bender, M.A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L.A., Lander, E.S., Dekker, J., 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326, 289–293. https://doi.org/10.1126/science.1181369

Liu, C., Li, Y., Semenov, M., Han, C., Baeg, G.H., Tan, Y., Zhang, Z., Lin, X., He, X., 2002. Control of beta-catenin phosphorylation/degradation by a dual-kinase mechanism. Cell 108, 837–847. https://doi.org/10.1016/s0092-8674(02)00685-2

Liu, P., Wakamiya, M., Shea, M.J., Albrecht, U., Behringer, R.R., Bradley, A., 1999. Requirement for Wnt3 in vertebrate axis formation. Nat Genet 22, 361–365. https://doi.org/10.1038/11932

Liu, Z., Yan, S., Wang, J., Xu, Y., Wang, Y., Zhang, S., Xu, X., Yang, Q., Zeng, X., Zhou, Y., Gu, X., Lu, S., Fu, Z., Fulton, D.J., Weintraub, N.L., Caldwell, R.B., Zhang, W., Wu, C., Liu, X.-L., Chen, J.-F., Ahmad, A., Kaddour-Djebbar, I., Al-Shabrawey, M., Li, Q., Jiang, X., Sun, Y., Sodhi, A., Smith, L., Hong, M., Huo, Y., 2017. Endothelial adenosine A2a receptor-mediated glycolysis is essential

for pathological retinal angiogenesis. Nat Commun 8, 584. https://doi.org/10.1038/s41467-017-00551-2

Logan, C.Y., Nusse, R., 2004. The Wnt Signaling Pathway in Development and Disease. Annual Review of Cell and Developmental Biology 20, 781–810. https://doi.org/10.1146/annurev.cellbio.20.010403.113126

López-Pérez, A.R., Balwierz, P.J., Lenhard, B., Muller, F., Wardle, F.C., Manfroid, I., Voz, M.L., Peers, B., 2021. Identification of downstream effectors of retinoic acid specifying the zebrafish pancreas by integrative genomics. Sci Rep 11, 22717. https://doi.org/10.1038/s41598-021-02039-y

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15, 550. https://doi.org/10.1186/s13059-014-0550-8

Lu, C.C., Robertson, E.J., 2004. Multiple roles for Nodal in the epiblast of the mouse embryo in the establishment of anterior-posterior patterning. Developmental Biology 273, 149–159. https://doi.org/10.1016/j.ydbio.2004.06.004

Lu, F., Liu, Y., Inoue, A., Suzuki, T., Zhao, K., Zhang, Y., 2016. Establishing chromatin regulatory landscape during mouse preimplantation development. Cell 165, 1375–1388. https://doi.org/10.1016/j.cell.2016.05.050

Lu, Y., Ren, X., Wang, Y., Han, J., 2018. Post-translational modifications and secretion of Wnt proteins. Biomedical Journal of Scientific & Technical Research 9.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., Richmond, T.J., 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 389, 251–260. https://doi.org/10.1038/38444

Lupiáñez, D.G., Spielmann, M., Mundlos, S., 2016. Breaking TADs: How Alterations of Chromatin Domains Result in Disease. Trends Genet 32, 225–237. https://doi.org/10.1016/j.tig.2016.01.003

Mallo, M., 2021. Of Necks, Trunks and Tails: Axial Skeletal Diversity among Vertebrates. Diversity 13, 289. https://doi.org/10.3390/d13070289

Mallo, M., 2018. Reassessing the Role of Hox Genes during Vertebrate Development and Evolution. Trends Genet 34, 209–217. https://doi.org/10.1016/j.tig.2017.11.007

Mallo, M., Wellik, D.M., Deschamps, J., 2010. Hox genes and regional patterning of the vertebrate body plan. Developmental Biology 344, 7–15. https://doi.org/10.1016/j.ydbio.2010.04.024

Mariani, L., Guo, X., Menezes, N.A., Drozd, A.M., Çakal, S.D., Wang, Q., Ferretti, E., 2021. A TALE/HOX code unlocks WNT signalling response towards paraxial mesoderm. Nat Commun 12, 5136. https://doi.org/10.1038/s41467-021-25370-4

Martin, B.L., Kimelman, D., 2010. Brachyury establishes the embryonic mesodermal progenitor niche. Genes Dev. 24, 2778–2783. https://doi.org/10.1101/gad.1962910

Matt, N., Dupé, V., Garnier, J.-M., Dennefeld, C., Chambon, P., Mark, M., Ghyselinck, N.B., 2005. Retinoic acid-dependent eye morphogenesis is orchestrated by

neural crest cells. Development 132, 4789–4800. https://doi.org/10.1242/dev.02031

Mbikay, M., Seidah, N.G., Chrétien, M., 2001. Neuroendocrine secretory protein 7B2: structure, expression and functions. Biochem J 357, 329–342.

McArthur, E., Capra, J.A., 2021. Topologically associating domain boundaries that are stable across diverse cell types are evolutionarily constrained and enriched for heritability. Am J Hum Genet 108, 269–283. https://doi.org/10.1016/j.ajhg.2021.01.001

McGinnis, W., Garber, R.L., Wirz, J., Kuroiwa, A., Gehring, W.J., 1984. A homologous protein-coding sequence in Drosophila homeotic genes and its conservation in other metazoans. Cell 37, 403–408. https://doi.org/10.1016/0092-8674(84)90370-2

McIntyre, D.C., Rakshit, S., Yallowitz, A.R., Loken, L., Jeannotte, L., Capecchi, M.R., Wellik, D.M., 2007. Hox patterning of the vertebrate rib cage. Development 134, 2981–2989. https://doi.org/10.1242/dev.007567

McMahon, A.P., Bradley, A., 1990. The Wnt-1 (int-1) proto-oncogene is required for development of a large region of the mouse brain. Cell 62, 1073–1085. https://doi.org/10.1016/0092-8674(90)90385-R

Mehta, S., Hingole, S., Chaudhary, V., 2021. The Emerging Mechanisms of Wnt Secretion and Signaling in Development. Frontiers in Cell and Developmental Biology 9.

Mi, H., Muruganujan, A., Ebert, D., Huang, X., Thomas, P.D., 2019. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res 47, D419–D426. https://doi.org/10.1093/nar/gky1038

Mifsud, B., Tavares-Cadete, F., Young, A.N., Sugar, R., Schoenfelder, S., Ferreira, L., Wingett, S.W., Andrews, S., Grey, W., Ewels, P.A., Herman, B., Happe, S., Higgs, A., LeProust, E., Follows, G.A., Fraser, P., Luscombe, N.M., Osborne, C.S., 2015. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. Nat Genet 47, 598–606. https://doi.org/10.1038/ng.3286

Migeotte, I., Omelchenko, T., Hall, A., Anderson, K.V., 2010. Rac1-Dependent Collective Cell Migration Is Required for Specification of the Anterior-Posterior Body Axis of the Mouse. PLOS Biology 8, e1000442. https://doi.org/10.1371/journal.pbio.1000442

Milewski, R.C., Chi, N.C., Li, J., Brown, C., Lu, M.M., Epstein, J.A., 2004. Identification of minimal enhancer elements sufficient for Pax3 expression in neural crest and implication of Tead2 as a regulator of Pax3. Development 131, 829–837. https://doi.org/10.1242/dev.00975

Mitiku, N., Baker, J.C., 2007. Genomic Analysis of Gastrulation and Organogenesis in the Mouse. Developmental Cell 13, 897–907. https://doi.org/10.1016/j.devcel.2007.10.004

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., Yamanaka, S., 2003. The homeoprotein Nanog is

required for maintenance of pluripotency in mouse epiblast and ES cells. Cell 113, 631–642. https://doi.org/10.1016/s0092-8674(03)00393-3

Molotkova, N., Molotkov, A., Sirbu, I.O., Duester, G., 2005. Requirement of mesodermal retinoic acid generated by Raldh2 for posterior neural transformation. Mech Dev 122, 145–155. https://doi.org/10.1016/j.mod.2004.10.008

Morimoto, M., Takahashi, Y., Endo, M., Saga, Y., 2005. The Mesp2 transcription factor establishes segmental borders by suppressing Notch activity. Nature 435, 354–359. https://doi.org/10.1038/nature03591

Müerköster, S., Arlt, A., Sipos, B., Witt, M., Großmann, M., Klöppel, G., Kalthoff, H., Fölsch, U.R., Schäfer, H., 2005. Increased Expression of the E3-Ubiquitin Ligase Receptor Subunit βTRCP1 Relates to Constitutive Nuclear Factor-κB Activation and Chemoresistance in Pancreatic Carcinoma Cells. Cancer Research 65, 1316–1324. https://doi.org/10.1158/0008-5472.CAN-04-1626

Murray, K., 1964. THE OCCURRENCE OF EPSILON-N-METHYL LYSINE IN HISTONES. Biochemistry 3, 10–15. https://doi.org/10.1021/bi00889a003

Narendra, V., Rocha, P.P., An, D., Raviram, R., Skok, J.A., Mazzoni, E.O., Reinberg, D., 2015. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. Science 347, 1017–1021. https://doi.org/10.1126/science.1262088

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H., Smith, A., 1998. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell 95, 379–391. https://doi.org/10.1016/s0092-8674(00)81769-9

Niederreither, K., McCaffery, P., Dräger, U.C., Chambon, P., Dollé, P., 1997. Restricted expression and retinoic acid-induced downregulation of the retinaldehyde dehydrogenase type 2 (RALDH-2) gene during mouse development. Mechanisms of Development 62, 67–78. https://doi.org/10.1016/S0925-4773(96)00653-3

Niederreither, K., Subbarayan, V., Dollé, P., Chambon, P., 1999. Embryonic retinoic acid synthesis is essential for early mouse post-implantation development. Nat Genet 21, 444–448. https://doi.org/10.1038/7788

Niederreither, K., Vermot, J., Messaddeq, N., Schuhbaur, B., Chambon, P., Dollé, P., 2001. Embryonic retinoic acid synthesis is essential for heart morphogenesis in the mouse. Development 128, 1019–1031. https://doi.org/10.1242/dev.128.7.1019

Nolte, M.J., Wang, Y., Deng, J.M., Swinton, P.G., Wei, C., Guindani, M., Schwartz, R.J., Behringer, R.R., 2014. Functional analysis of limb transcriptional enhancers in the mouse. Evol Dev 16, 207–223. https://doi.org/10.1111/ede.12084

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., Gribnau, J., Barillot, E., Blüthgen, N., Dekker, J., Heard, E., 2012. Spatial partitioning of the regulatory landscape of

the X-inactivation centre. Nature 485, 381–385. https://doi.org/10.1038/nature11049

Novitch, B.G., Wichterle, H., Jessell, T.M., Sockanathan, S., 2003. A requirement for retinoic acid-mediated transcriptional activation in ventral neural patterning and motor neuron specification. Neuron 40, 81–95. https://doi.org/10.1016/j.neuron.2003.08.006

Nowotschin, S., Costello, I., Piliszek, A., Kwon, G.S., Mao, C., Klein, W.H., Robertson, E.J., Hadjantonakis, A.-K., 2013. The T-box transcription factor Eomesodermin is essential for AVE induction in the mouse embryo. Genes Dev. 27, 997–1002. https://doi.org/10.1101/gad.215152.113

Oishi, I., Suzuki, H., Onishi, N., Takada, R., Kani, S., Ohkawara, B., Koshida, I., Suzuki, K., Yamada, G., Schwabe, G.C., Mundlos, S., Shibuya, H., Takada, S., Minami, Y., 2003. The receptor tyrosine kinase Ror2 is involved in non-canonical Wnt5a/JNK signalling pathway. Genes Cells 8, 645–654. https://doi.org/10.1046/j.1365-2443.2003.00662.x

Okada, Y., Takeda, S., Tanaka, Y., Belmonte, J.-C.I., Hirokawa, N., 2005. Mechanism of Nodal Flow: A Conserved Symmetry Breaking Event in Left-Right Axis Determination. Cell 121, 633–644. https://doi.org/10.1016/j.cell.2005.04.008

Okazawa, H., Okamoto, K., Ishino, F., Ishino-Kaneko, T., Takeda, S., Toyoda, Y., Muramatsu, M., Hamada, H., 1991. The oct3 gene, a gene for an embryonic transcription factor, is controlled by a retinoic acid repressible enhancer. EMBO J 10, 2997–3005. https://doi.org/10.1002/j.1460-2075.1991.tb07850.x

Olivera-Martinez, I., Storey, K.G., 2007. Wnt signals provide a timing mechanism for the FGF-retinoid differentiation switch during vertebrate body axis extension. Development 134, 2125–2135. https://doi.org/10.1242/dev.000216

Osterwalder, M., Barozzi, I., Tissières, V., Fukuda-Yuzawa, Y., Mannion, B.J., Afzal, S.Y., Lee, E.A., Zhu, Y., Plajzer-Frick, I., Pickle, C.S., Kato, M., Garvin, T.H., Pham, Q.T., Harrington, A.N., Akiyama, J.A., Afzal, V., Lopez-Rios, J., Dickel, D.E., Visel, A., Pennacchio, L.A., 2018. Enhancer redundancy provides phenotypic robustness in mammalian development. Nature 554, 239–243. https://doi.org/10.1038/nature25461

Ozeki, A., Tsukamoto, I., 1999. Retinoic acid repressed the expression of c-fos and c-jun and induced apoptosis in regenerating rat liver after partial hepatectomy. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1450, 308–319. https://doi.org/10.1016/S0167-4889(99)00063-4

Packer, A.I., Crotty, D.A., Elwell, V.A., Wolgemuth, D.J., 1998. Expression of the murine Hoxa4 gene requires both autoregulation and a conserved retinoic acid response element. Development 125, 1991–1998. https://doi.org/10.1242/dev.125.11.1991

Paik, W.K., Kim, S., 1967. E-N-dimethyllysine in histones. Biochemical and Biophysical Research Communications 27, 479–483. https://doi.org/10.1016/S0006-291X(67)80010-X

Palmeirim, I., Henrique, D., Ish-Horowicz, D., Pourquié, O., 1997. Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and

somitogenesis. Cell 91, 639–648. https://doi.org/10.1016/s0092-8674(00)80451-1

Panáková, D., Sprong, H., Marois, E., Thiele, C., Eaton, S., 2005. Lipoprotein particles are required for Hedgehog and Wingless signalling. Nature 435, 58–65. https://doi.org/10.1038/nature03504

Pandur, P., Läsche, M., Eisenberg, L.M., Kühl, M., 2002. Wnt-11 activation of a non-canonical Wnt signalling pathway is required for cardiogenesis. Nature 418, 636–641. https://doi.org/10.1038/nature00921

Papa, F., Lippolis, R., Sardaro, N., Gnoni, A., Scacco, S., 2017. All trans retinoic acid depresses the content and activity of the mitochondrial ATP synthase in human keratinocytes. Biochemical and Biophysical Research Communications 482, 301–304. https://doi.org/10.1016/j.bbrc.2016.11.058

Perea-Gomez, A., Vella, F.D.J., Shawlot, W., Oulad-Abdelghani, M., Chazaud, C., Meno, C., Pfister, V., Chen, L., Robertson, E., Hamada, H., Behringer, R.R., Ang, S.-L., 2002. Nodal Antagonists in the Anterior Visceral Endoderm Prevent the Formation of Multiple Primitive Streaks. Developmental Cell 3, 745–756. https://doi.org/10.1016/S1534-5807(02)00321-0

Perry, M.W., Boettiger, A.N., Bothma, J.P., Levine, M., 2010. Shadow enhancers foster robustness of Drosophila gastrulation. Curr Biol 20, 1562–1567. https://doi.org/10.1016/j.cub.2010.07.043

Perry, M.W., Boettiger, A.N., Levine, M., 2011. Multiple enhancers ensure precision of gap gene-expression patterns in the Drosophila embryo. Proc Natl Acad Sci U S A 108, 13570–13575. https://doi.org/10.1073/pnas.1109873108

Peters, A.H.F.M., Kubicek, S., Mechtler, K., O'Sullivan, R.J., Derijck, A.A.H.A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y., Martens, J.H.A., Jenuwein, T., 2003. Partitioning and Plasticity of Repressive Histone Methylation States in Mammalian Chromatin. Molecular Cell. https://doi.org/10.1016/S1097-2765(03)00477-5

Phillips-Cremins, J.E., Sauria, M.E.G., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S.K., Ong, C.-T., Hookway, T.A., Guo, C., Sun, Y., Bland, M.J., Wagstaff, W., Dalton, S., McDevitt, T.C., Sen, R., Dekker, J., Taylor, J., Corces, V.G., 2013. Architectural Protein Subclasses Shape 3D Organization of Genomes during Lineage Commitment. Cell 153, 1281–1295. https://doi.org/10.1016/j.cell.2013.04.053

Picelli, S., Faridani, O.R., Björklund, Å.K., Winberg, G., Sagasser, S., Sandberg, R., 2014. Full-length RNA-seq from single cells using Smart-seq2. Nat Protoc 9, 171–181. https://doi.org/10.1038/nprot.2014.006

Pijuan-Sala, B., Griffiths, J.A., Guibentif, C., Hiscock, T.W., Jawaid, W., Calero-Nieto, F.J., Mulas, C., Ibarra-Soria, X., Tyser, R.C.V., Ho, D.L.L., Reik, W., Srinivas, S., Simons, B.D., Nichols, J., Marioni, J.C., Göttgens, B., 2019. A single-cell molecular map of mouse gastrulation and early organogenesis. Nature 566, 490–495. https://doi.org/10.1038/s41586-019-0933-9

Pikarsky, E., Sharir, H., Ben-Shushan, E., Bergman, Y., 1994. Retinoic acid represses Oct-3/4 gene expression through several retinoic acid-responsive elements located in the promoter-enhancer region. Mol Cell Biol 14, 1026–1038.

Pinson, K.I., Brennan, J., Monkley, S., Avery, B.J., Skarnes, W.C., 2000. An LDL-receptor-related protein mediates Wnt signalling in mice. Nature 407, 535–538. https://doi.org/10.1038/35035124

Pollard, K.S., Hubisz, M.J., Rosenbloom, K.R., Siepel, A., 2010. Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res 20, 110–121. https://doi.org/10.1101/gr.097857.109

Qian, L., Mahaffey, J.P., Alcorn, H.L., Anderson, K.V., 2011. Tissue-specific roles of Axin2 in the inhibition and activation of Wnt signaling in the mouse embryo. Proceedings of the National Academy of Sciences 108, 8692–8697. https://doi.org/10.1073/pnas.1100328108

Quinlan, G.A., Williams, E.A., Tan, S.S., Tam, P.P., 1995. Neuroectodermal fate of epiblast cells in the distal region of the mouse egg cylinder: implication for body plan organization during early embryogenesis. Development 121, 87–98. https://doi.org/10.1242/dev.121.1.87

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., Wysocka, J., 2011. A unique chromatin signature uncovers early developmental enhancers in humans. Nature. https://doi.org/10.1038/nature09692

Rao, D.M., Shackleford, M.T., Bordeaux, E.K., Sottnik, J.L., Ferguson, R.L., Yamamoto, T.M., Wellberg, E.A., Bitler, B.G., Sikora, M.J., 2019. Wnt family member 4 (WNT4) and WNT3A activate cell-autonomous Wnt signaling independent of porcupine O-acyltransferase or Wnt secretion. J Biol Chem 294, 19950–19966. https://doi.org/10.1074/jbc.RA119.009615

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., Aiden, E.L., 2014. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell 159, 1665–1680. https://doi.org/10.1016/j.cell.2014.11.021

Raya, Á., Kawakami, Y., Rodríguez-Esteban, C., Ibañes, M., Rasskin-Gutman, D., Rodríguez-León, J., Büscher, D., Feijó, J.A., Izpisúa Belmonte, J.C., 2004. Notch activity acts as a sensor for extracellular calcium during vertebrate left–right determination. Nature 427, 121–128. https://doi.org/10.1038/nature02190

Ribes, V., Stutzmann, F., Bianchetti, L., Guillemot, F., Dollé, P., Le Roux, I., 2008. Combinatorial signalling controls Neurogenin2 expression at the onset of spinal neurogenesis. Developmental Biology 321, 470–481. https://doi.org/10.1016/j.ydbio.2008.06.003

Rim, E.Y., Clevers, H., Nusse, R., 2022. The Wnt Pathway: From Signaling Mechanisms to Synthetic Modulators. Annu Rev Biochem 91, 571–598. https://doi.org/10.1146/annurev-biochem-040320-103615

Rios-Esteves, J., Haugen, B., Resh, M.D., 2014. Identification of key residues and regions important for porcupine-mediated Wnt acylation. J Biol Chem 289, 17009–17019. https://doi.org/10.1074/jbc.M114.561209

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., Mesirov, J.P., 2011. Integrative genomics viewer. Nat Biotechnol 29, 24–26. https://doi.org/10.1038/nbt.1754

Rosa, A., Brivanlou, A.H., 2011. A regulatory circuitry comprised of miR-302 and the transcription factors OCT4 and NR2F2 regulates human embryonic stem cell differentiation. EMBO J 30, 237–248. https://doi.org/10.1038/emboj.2010.319

Sabo, P.J., Kuehn, M.S., Thurman, R., Johnson, B.E., Johnson, E.M., Cao, H., Yu, M., Rosenzweig, E., Goldy, J., Haydock, A., Weaver, M., Shafer, A., Lee, K., Neri, F., Humbert, R., Singer, M.A., Richmond, T.A., Dorschner, M.O., McArthur, M., Hawrylycz, M., Green, R.D., Navas, P.A., Noble, W.S., Stamatoyannopoulos, J.A., 2006. Genome-scale mapping of DNase I sensitivity in vivo using tiling DNA microarrays. Nat Methods 3, 511–518. https://doi.org/10.1038/nmeth890

Saga, Y., Takeda, H., 2001. The making of the somite: molecular events in vertebrate segmentation. Nat Rev Genet 2, 835–845. https://doi.org/10.1038/35098552

Sakai, Y., Meno, C., Fujii, H., Nishino, J., Shiratori, H., Saijoh, Y., Rossant, J., Hamada, H., 2001. The retinoic acid-inactivating enzyme CYP26 is essential for establishing an uneven distribution of retinoic acid along the anterio-posterior axis within the mouse embryo. Genes Dev. 15, 213–225. https://doi.org/10.1101/gad.851501

Sandell, L.L., Lynn, M.L., Inman, K.E., McDowell, W., Trainor, P.A., 2012. RDH10 oxidation of Vitamin A is a critical control step in synthesis of retinoic acid during mouse embryogenesis. PLoS One 7, e30698. https://doi.org/10.1371/journal.pone.0030698

Sansom, S.N., Griffiths, D.S., Faedo, A., Kleinjan, D.-J., Ruan, Y., Smith, J., Heyningen, V. van, Rubenstein, J.L., Livesey, F.J., 2009. The Level of the Transcription Factor Pax6 Is Essential for Controlling the Balance between Neural Stem Cell Self-Renewal and Neurogenesis. PLOS Genetics 5, e1000511. https://doi.org/10.1371/journal.pgen.1000511

Satoh, W., Gotoh, T., Tsunematsu, Y., Aizawa, S., Shimono, A., 2006. Sfrp1 and Sfrp2 regulate anteroposterior axis elongation and somite segmentation during mouse embryogenesis. Development 133, 989–999. https://doi.org/10.1242/dev.02274

Savory, J.G.A., Mansfield, M., Rijli, F.M., Lohnes, D., 2011. Cdx mediates neural tube closure through transcriptional regulation of the planar cell polarity gene Ptk7. Development 138, 1361–1370. https://doi.org/10.1242/dev.056622

Sawado, T., Halow, J., Bender, M.A., Groudine, M., 2003. The β-globin locus control region (LCR) functions primarily by enhancing the transition from transcription initiation to elongation. Genes Dev. 17, 1009–1018. https://doi.org/10.1101/gad.1072303

Schiffrin, E.L., 2007. The Endothelin System, in: Lip, G.Y.H., Hall, J.E. (Eds.), Comprehensive Hypertension. Mosby, Philadelphia, pp. 317–323. https://doi.org/10.1016/B978-0-323-03961-1.50031-3

Schleif, C., Gewiss, R., Griswold, M., 2023. Chromatin Remodeling via Retinoic Acid Action during Murine Spermatogonial Development. Life 13, 690. https://doi.org/10.3390/life13030690

Schmidl, C., Rendeiro, A.F., Sheffield, N.C., Bock, C., 2015. ChIPmentation: fast, robust, low-input ChIP-seq for histones and transcription factors. Nat Methods 12, 963–965. https://doi.org/10.1038/nmeth.3542

Schones, D.E., Cui, K., Cuddapah, S., Roh, T.-Y., Barski, A., Wang, Z., Wei, G., Zhao, K., 2008. Dynamic regulation of nucleosome positioning in the human genome. Cell 132, 887–898. https://doi.org/10.1016/j.cell.2008.02.022

Schröter, C., Herrgen, L., Cardona, A., Brouhard, G.J., Feldman, B., Oates, A.C., 2008. Dynamics of zebrafish somitogenesis. Dev Dyn 237, 545–553. https://doi.org/10.1002/dvdy.21458

Schwab, A., Fabian, A., Hanley, P.J., Stock, C., 2012. Role of ion channels and transporters in cell migration. Physiol Rev 92, 1865–1913. https://doi.org/10.1152/physrev.00018.2011

Seo, S., Kim, Y.A., Lee, J., Lee, Seunghwan, Kim, J., Lee, Seunghee, 2022. Fat3 regulates neural progenitor cells by promoting Yap activity during spinal cord development. Sci Rep 12, 14726. https://doi.org/10.1038/s41598-022-19029-3

Sibbritt, T., Ip, C.K., Khoo, P.-L., Wilkie, E., Jones, V., Sun, J.Q.J., Shen, J.X., Peng, G., Han, J.-D.J., Jing, N., Osteil, P., Ramialison, M., Tam, P.P.L., Fossat, N., 2018. A gene regulatory network anchored by LIM homeobox 1 for embryonic head development. Genesis 56, e23246. https://doi.org/10.1002/dvg.23246

Siegfried, E., Chou, T.B., Perrimon, N., 1992. wingless signaling acts through zeste-white 3, the Drosophila homolog of glycogen synthase kinase-3, to regulate engrailed and establish cell fate. Cell 71, 1167–1179. https://doi.org/10.1016/s0092-8674(05)80065-0

Simas, T., Correia, R.B., Rocha, L.M., 2021. The distance backbone of complex networks. Journal of Complex Networks 9, cnab021. https://doi.org/10.1093/comnet/cnab021

Singh, N.P., De Kumar, B., Paulson, A., Parrish, M.E., Zhang, Y., Florens, L., Conaway, J.W., Si, K., Krumlauf, R., 2020. A six-amino-acid motif is a major determinant in functional evolution of HOX1 proteins. Genes Dev 34, 1680–1696. https://doi.org/10.1101/gad.342329.120

Slusarski, Corces, V.G., Moon, R.T., 1997. Interaction of Wnt and a Frizzled homologue triggers G-protein-linked phosphatidylinositol signalling. Nature 390, 410–413. https://doi.org/10.1038/37138

Slusarski, D.C., Yang-Snyder, J., Busa, W.B., Moon, R.T., 1997. Modulation of embryonic intracellular Ca2+ signaling by Wnt-5A. Dev Biol 182, 114–120. https://doi.org/10.1006/dbio.1996.8463

Soshnikova, N., Duboule, D., 2009. Epigenetic regulation of vertebrate Hox genes: A dynamic equilibrium. Epigenetics. https://doi.org/10.4161/epi.4.8.10132

Srinivas, S., Rodriguez, T., Clements, M., Smith, J.C., Beddington, R.S.P., 2004. Active cell migration drives the unilateral movements of the anterior visceral endoderm. Development 131, 1157–1164. https://doi.org/10.1242/dev.01005

Stafford, D.A., Monica, S.D., Harland, R.M., 2014. Follistatin interacts with Noggin in the development of the axial skeleton. Mechanisms of Development 131, 78–85. https://doi.org/10.1016/j.mod.2013.10.001

Staveley, B., 2018. Molecular and Developmental Biology. Vertebrate Development II. Axes and Germ Layers (BIOL3530).

Stern, C.D., Charite, J., Deschamps, J., Duboule, D., Durston, A.J., Kmita, M., Nicolas, J.-F., Palmeirim, I., Smith, J.C., Wolpert, L., 2006. Head-tail patterning of the vertebrate embryo: one, two or many unresolved problems? Int. J. Dev. Biol. 50, 3–15. https://doi.org/10.1387/ijdb.052095cs

Steventon, B., Martinez-Arias, A., 2017. Evo-engineering and the cellular and molecular origins of the vertebrate spinal cord. Dev Biol 432, 3–13. https://doi.org/10.1016/j.ydbio.2017.01.021

Strutt, D.I., Weber, U., Mlodzik, M., 1997. The role of RhoA in tissue polarity and Frizzled signalling. Nature 387, 292–295. https://doi.org/10.1038/387292a0

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., Satija, R., 2019. Comprehensive Integration of Single-Cell Data. Cell 177, 1888-1902.e21. https://doi.org/10.1016/j.cell.2019.05.031

Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., Jensen, L.J., Mering, C. von, 2019. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 47, D607–D613. https://doi.org/10.1093/nar/gky1131

Tahara, N., Kawakami, H., Chen, K.Q., Anderson, A., Yamashita Peterson, M., Gong, W., Shah, P., Hayashi, S., Nishinakamura, R., Nakagawa, Y., Garry, D.J., Kawakami, Y., 2019. Sall4 regulates neuromesodermal progenitors and their descendants during body elongation in mouse embryos. Development 146, dev177659. https://doi.org/10.1242/dev.177659

Takada, R., Satomi, Y., Kurata, T., Ueno, N., Norioka, S., Kondoh, H., Takao, T., Takada, S., 2006. Monounsaturated Fatty Acid Modification of Wnt Protein: Its Role in Wnt Secretion. Developmental Cell 11, 791–801. https://doi.org/10.1016/j.devcel.2006.10.003

Takada, S., Stark, K.L., Shea, M.J., Vassileva, G., McMahon, J.A., McMahon, A.P., 1994. Wnt-3a regulates somite and tailbud formation in the mouse embryo. Genes Dev. 8, 174–189. https://doi.org/10.1101/gad.8.2.174

Takahashi, J., Ohbayashi, A., Oginuma, M., Saito, D., Mochizuki, A., Saga, Y., Takada, S., 2010. Analysis of Ripply1/2-deficient mouse embryos reveals a mechanism underlying the rostro-caudal patterning within a somite. Developmental Biology 342, 134–145. https://doi.org/10.1016/j.ydbio.2010.03.015

Takaoka, K., Yamamoto, M., Hamada, H., 2011. Origin and role of distal visceral endoderm, a group of cells that determines anterior–posterior polarity of the mouse embryo. Nat Cell Biol 13, 743–752. https://doi.org/10.1038/ncb2251

Tam, P.P., 1981. The control of somitogenesis in mouse embryos. J Embryol Exp Morphol 65 Suppl, 103–128.

Tam, P.P., Beddington, R.S., 1987. The formation of mesodermal tissues in the mouse embryo during gastrulation and early organogenesis. Development 99, 109–126. https://doi.org/10.1242/dev.99.1.109

Tam, P.P., Behringer, R.R., 1997. Mouse gastrulation: the formation of a mammalian body plan. Mech Dev 68, 3–25. https://doi.org/10.1016/s0925-4773(97)00123-8

Tam, P.P., Parameswaran, M., Kinder, S.J., Weinberger, R.P., 1997. The allocation of epiblast cells to the embryonic heart and other mesodermal lineages: the role of ingression and tissue movement during gastrulation. Development 124, 1631–1642. https://doi.org/10.1242/dev.124.9.1631

Tam, P.P.L., Khoo, P.-L., Lewis, S.L., Bildsoe, H., Wong, N., Tsang, T.E., Gad, J.M., Robb, L., 2007. Sequential allocation and global pattern of movement of the definitive endoderm in the mouse embryo during gastrulation. Development 134, 251–260. https://doi.org/10.1242/dev.02724

Tanaka, Y., Okada, Y., Hirokawa, N., 2005. FGF-induced vesicular release of Sonic hedgehog and retinoic acid in leftward nodal flow is critical for left–right determination. Nature 435, 172–177. https://doi.org/10.1038/nature03494

Tekko, T., Lozovska, A., Nóvoa, A., Mallo, M., 2022. Assessing Myf5 and Lbx1 contribution to carapace development by reproducing their turtle-specific signatures in mouse embryos. Dev Dyn. https://doi.org/10.1002/dvdy.502

The ENCODE Project Consortium, Moore, J.E., Purcaro, M.J., Pratt, H.E., Epstein, C.B., Shoresh, N., Adrian, J., Kawli, T., Davis, C.A., Dobin, A., Kaul, R., Halow, J., Van Nostrand, E.L., Freese, P., Gorkin, D.U., Shen, Y., He, Y., Mackiewicz, M., Pauli-Behn, F., Williams, B.A., Mortazavi, A., Keller, C.A., Zhang, X.-O., Elhajjajy, S.I., Huey, J., Dickel, D.E., Snetkova, V., Wei, X., Wang, X., Rivera-Mulia, J.C., Rozowsky, J., Zhang, Jing, Chhetri, S.B., Zhang, Jialing, Victorsen, A., White, K.P., Visel, A., Yeo, G.W., Burge, C.B., Lécuyer, E., Gilbert, D.M., Dekker, J., Rinn, J., Mendenhall, E.M., Ecker, J.R., Kellis, M., Klein, R.J., Noble, W.S., Kundaje, A., Guigó, R., Farnham, P.J., Cherry, J.M., Myers, R.M., Ren, B., Graveley, B.R., Gerstein, M.B., Pennacchio, L.A., Snyder, M.P., Bernstein, B.E., Wold, B., Hardison, R.C., Gingeras, T.R., Stamatoyannopoulos, J.A., Weng, Z., 2020. Expanded encyclopaedias of DNA elements in the human and mouse genomes. Nature 583, 699–710. https://doi.org/10.1038/s41586-020-2493-4

Tiana, M., Lopez-Jimenez, E., de Aja, J.S., Barral, A., Victorino, J., Badia-Careaga, C., Rollan, I., Rouco, R., Santos, E., Sanchez-Iranzo, H., Acemel, R.D., Torroja, C., Adan, J., Andres-Leon, E., Gomez-Skarmeta, J.L., Giovinazzo, G., Sanchez-Cabo, F., Manzanares, M., 2022. Pluripotency factors regulate the

onset of Hox cluster activation in the early embryo. Sci Adv 8, eabo3583. https://doi.org/10.1126/sciadv.abo3583

Tourniaire, F., Musinovic, H., Gouranton, E., Astier, J., Marcotorchino, J., Arreguin, A., Bernot, D., Palou, A., Bonet, M.L., Ribot, J., Landrier, J.-F., 2015. All-trans retinoic acid induces oxidative phosphorylation and mitochondria biogenesis in adipocytes. J Lipid Res 56, 1100–1109. https://doi.org/10.1194/jlr.M053652

Turek, V.F., Trevaskis, J.L., Levin, B.E., Dunn-Meynell, A.A., Irani, B., Gu, G., Wittmer, C., Griffin, P.S., Vu, C., Parkes, D.G., Roth, J.D., 2010. Mechanisms of Amylin/Leptin Synergy in Rodent Models. Endocrinology 151, 143–152. https://doi.org/10.1210/en.2009-0546

Tzouanacou, E., Wegener, A., Wymeersch, F.J., Wilson, V., Nicolas, J.-F., 2009. Redefining the progression of lineage segregations during mammalian embryogenesis by clonal analysis. Dev Cell 17, 365–376. https://doi.org/10.1016/j.devcel.2009.08.002

van den Akker, E., Fromental-Ramain, C., de Graaff, W., Le Mouellic, H., Brûlet, P., Chambon, P., Deschamps, J., 2001. Axial skeletal patterning in mice lacking all paralogous group 8 Hox genes. Development 128, 1911–1921. https://doi.org/10.1242/dev.128.10.1911

Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., Jaenisch, R., 2013. One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. Cell 153, 910–918. https://doi.org/10.1016/j.cell.2013.04.025

Wang, Q., Symes, A.J., Kane, C.A., Freeman, A., Nariculam, J., Munson, P., Thrasivoulou, C., Masters, J.R.W., Ahmed, A., 2010. A Novel Role for Wnt/Ca2+ Signaling in Actin Cytoskeleton Remodeling and Cell Motility in Prostate Cancer. PLOS ONE 5, e10456. https://doi.org/10.1371/journal.pone.0010456

Wang, Y., Song, F., Zhang, B., Zhang, L., Xu, J., Kuang, D., Li, D., Choudhary, M.N.K., Li, Y., Hu, M., Hardison, R., Wang, T., Yue, F., 2018. The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. Genome Biology 19, 151. https://doi.org/10.1186/s13059-018-1519-9

Wang, Z., Zang, C., Cui, K., Schones, D.E., Barski, A., Peng, W., Zhao, K., 2009. Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. Cell 138, 1019–1031. https://doi.org/10.1016/j.cell.2009.06.049

Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Peng, W., Zhang, M.Q., Zhao, K., 2008. Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet 40, 897–903. https://doi.org/10.1038/ng.154

Waymack, R., Fletcher, A., Enciso, G., Wunderlich, Z., 2020. Shadow enhancers can suppress input transcription factor noise through distinct regulatory logic. eLife 9, e59351. https://doi.org/10.7554/eLife.59351

Wellik, D.M., Capecchi, M.R., 2003. Hox10 and Hox11 genes are required to globally pattern the mammalian skeleton. Science 301, 363–367. https://doi.org/10.1126/science.1085672

Wetering, M. van de, Cavallo, R., Dooijes, D., Beest, M. van, Es, J. van, Loureiro, J., Ypma, A., Hursh, D., Jones, T., Bejsovec, A., Peifer, M., Mortin, M., Clevers, H., 1997. Armadillo Coactivates Transcription Driven by the Product of the Drosophila Segment Polarity Gene dTCF. Cell 88, 789–799. https://doi.org/10.1016/S0092-8674(00)81925-X

Widom, J., 1998. Structure, dynamics, and function of chromatin in vitro. Annual Review of Biophysics and Biomolecular Structure. https://doi.org/10.1146/annurev.biophys.27.1.285

Widschwendter, M., Widschwendter, A., Welte, T., Daxenbichler, G., Zeimet, A.G., Bergant, A., Berger, J., Peyrat, J.P., Michel, S., Doppler, W., Marth, C., 1999. Retinoic acid modulates prolactin receptor expression and prolactin-induced STAT-5 activation in breast cancer cells in vitro. Br J Cancer 79, 204–210. https://doi.org/10.1038/sj.bjc.6690034

William, D.A., Saitta, B., Gibson, J.D., Traas, J., Markov, V., Gonzalez, D.M., Sewell, W., Anderson, D.M., Pratt, S.C., Rappaport, E.F., Kusumi, K., 2007. Identification of oscillatory genes in somitogenesis from functional genomic analysis of a human mesenchymal stem cell model. Developmental Biology 305, 172–186. https://doi.org/10.1016/j.ydbio.2007.02.007

Wilson, L., Gale, E., Maden, M., 2003. The role of retinoic acid in the morphogenesis of the neural tube. Journal of Anatomy 203, 357–368. https://doi.org/10.1046/j.1469-7580.2003.00230.x

Wilson, V., Beddington, R.S., 1996. Cell fate and morphogenetic movement in the late mouse primitive streak. Mech Dev 55, 79–89. https://doi.org/10.1016/0925-4773(95)00493-9

Wilson, V., Olivera-Martinez, I., Storey, K.G., 2009. Stem cells, signals and vertebrate body axis extension. Development 136, 1591–1604. https://doi.org/10.1242/dev.021246

Winnier, G., Blessing, M., Labosky, P.A., Hogan, B.L., 1995. Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. Genes Dev. 9, 2105–2116. https://doi.org/10.1101/gad.9.17.2105

Winston, J.T., Strack, P., Beer-Romero, P., Chu, C.Y., Elledge, S.J., Harper, J.W., 1999. The SCFβ-TRCP–ubiquitin ligase complex associates specifically with phosphorylated destruction motifs in IκBα and β-catenin and stimulates IκBα ubiquitination in vitro. Genes Dev. 13, 270–283.

WINTERS, A.J., COLSTON, C., MACDONALD, P.C., PORTER, J.C., 1975. Fetal Plasma Prolactin Levels1. The Journal of Clinical Endocrinology & Metabolism 41, 626–629. https://doi.org/10.1210/jcem-41-3-626

Woltering, J.M., Noordermeer, D., Leleu, M., Duboule, D., 2014. Conservation and Divergence of Regulatory Strategies at Hox Loci and the Origin of Tetrapod Digits. PLOS Biology 12, e1001773. https://doi.org/10.1371/journal.pbio.1001773

Wong, C., Mahapatra, N.R., Chitbangonsyn, S., Mahboubi, P., Mahata, M., Mahata, S.K., O'Connor, D.T., 2003. The angiotensin II receptor (Agtr1a): functional regulatory polymorphisms in a locus genetically linked to blood pressure variation in the mouse. Physiological Genomics 14, 83–93. https://doi.org/10.1152/physiolgenomics.00162.2002

Wunderlich, Z., Bragdon, M.D.J., Vincent, B.J., White, J.A., Estrada, J., DePace, A.H., 2015. Krüppel Expression Levels Are Maintained through Compensatory Evolution of Shadow Enhancers. Cell Reports 12, 1740–1747. https://doi.org/10.1016/j.celrep.2015.08.021

Wymeersch, F.J., Huang, Y., Blin, G., Cambray, N., Wilkie, R., Wong, F.C., Wilson, V., 2016. Position-dependent plasticity of distinct progenitor types in the primitive streak. eLife 5, e10042. https://doi.org/10.7554/eLife.10042

Wymeersch, F.J., Wilson, V., Tsakiridis, A., 2021. Understanding axial progenitor biology in vivo and in vitro. Development 148, dev180612. https://doi.org/10.1242/dev.180612

Yamaguchi, T.P., Bradley, A., McMahon, A.P., Jones, S., 1999. A Wnt5a pathway underlies outgrowth of multiple structures in the vertebrate embryo. Development 126, 1211–1223. https://doi.org/10.1242/dev.126.6.1211

Yamamoto, M., Beppu, H., Takaoka, K., Meno, C., Li, E., Miyazono, K., Hamada, H., 2009. Antagonism between Smad1 and Smad2 signaling determines the site of distal visceral endoderm formation in the mouse embryo. Journal of Cell Biology 184, 323–334. https://doi.org/10.1083/jcb.200808044

Yamamoto, M., Saijoh, Y., Perea-Gomez, A., Shawlot, W., Behringer, R.R., Ang, S.-L., Hamada, H., Meno, C., 2004. Nodal antagonists regulate formation of the anteroposterior axis of the mouse embryo. Nature 428, 387–392. https://doi.org/10.1038/nature02418

Yanagisawa, H., Clouthier, D.E., Richardson, J.A., Charité, J., Olson, E.N., 2003. Targeted deletion of a branchial arch-specific enhancer reveals a role of dHAND in craniofacial development. Development 130, 1069–1078. https://doi.org/10.1242/dev.00337

Yeung, R.K., Xiang, Z.-H., Tsang, S.-Y., Li, R., Ho, T.Y.C., Li, Q., Hui, C.-K., Sham, P.-C., Qiao, M.-Q., Xue, H., 2018. Gabrb2-knockout mice displayed schizophrenia-like and comorbid phenotypes with interneuron–astrocyte–microglia dysregulation. Transl Psychiatry 8, 1–14. https://doi.org/10.1038/s41398-018-0176-9

Yoon, J.K., Wold, B., 2000. The bHLH regulator pMesogenin1 is required for maturation and segmentation of paraxial mesoderm. Genes Dev 14, 3204–3214. https://doi.org/10.1101/gad.850000

Yoshikawa, S., McKinnon, R.D., Kokel, M., Thomas, J.B., 2003. Wnt-mediated axon guidance via the Drosophila Derailed receptor. Nature 422, 583–588. https://doi.org/10.1038/nature01522

Yoshikawa, Y., Fujimori, T., McMahon, A.P., Takada, S., 1997. Evidence that absence of Wnt-3a signaling promotes neuralization instead of paraxial mesoderm

development in the mouse. Dev Biol 183, 234–242. https://doi.org/10.1006/dbio.1997.8502

Young, T., Rowland, J.E., van de Ven, C., Bialecka, M., Novoa, A., Carapuco, M., van Nes, J., de Graaff, W., Duluc, I., Freund, J.-N., Beck, F., Mallo, M., Deschamps, J., 2009. Cdx and Hox genes differentially regulate posterior axial growth in mammalian embryos. Dev Cell 17, 516–526. https://doi.org/10.1016/j.devcel.2009.08.010

Zelent, A., Krust, A., Petkovich, M., Kastner, P., Chambon, P., 1989. Cloning of murine α and β retinoic acid receptors and a novel receptor γ predominantly expressed in skin. Nature 339, 714–717. https://doi.org/10.1038/339714a0

Zhang, R., Kang, R., Klionsky, D.J., Tang, D., 2022. Ion Channels and Transporters in Autophagy. Autophagy 18, 4–23. https://doi.org/10.1080/15548627.2021.1885147

Zhang, X.M., Ramalho-Santos, M., McMahon, A.P., 2001. Smoothened mutants reveal redundant roles for Shh and Ihh signaling including regulation of L/R asymmetry by the mouse node. Cell 105, 781–792.

Zhao, R., Watt, A.J., Battle, M.A., Li, J., Bondow, B.J., Duncan, S.A., 2008. Loss of both GATA4 and GATA6 blocks cardiac myocyte differentiation and results in acardia in mice. Dev Biol 317, 614–619. https://doi.org/10.1016/j.ydbio.2008.03.013

Zhao, W., Ajima, R., Ninomiya, Y., Saga, Y., 2015. Segmental border is defined by Ripply2-mediated Tbx6 repression independent of Mesp2. Developmental Biology 400, 105–117. https://doi.org/10.1016/j.ydbio.2015.01.020

Zhao, X., Sirbu, I.O., Mic, F.A., Molotkova, N., Molotkov, A., Kumar, S., Duester, G., 2009. Retinoic acid promotes limb induction through effects on body axis extension but is unnecessary for limb patterning. Curr Biol 19, 1050–1057. https://doi.org/10.1016/j.cub.2009.04.059

Zhao, Y., Vuckovic, M., Yoo, H.S., Fox, N., Rodriguez, A., McKessy, K., Napoli, J.L., 2021. Retinoic acid exerts sexually dimorphic effects on muscle energy metabolism and function. J Biol Chem 297, 101101. https://doi.org/10.1016/j.jbc.2021.101101

Zheng, N., Shabek, N., 2017. Ubiquitin Ligases: Structure, Function, and Regulation. Annu Rev Biochem 86, 129–157. https://doi.org/10.1146/annurev-biochem-060815-014922

Zuin, J., Dixon, J.R., van der Reijden, M.I.J.A., Ye, Z., Kolovos, P., Brouwer, R.W.W., van de Corput, M.P.C., van de Werken, H.J.G., Knoch, T.A., van IJcken, W.F.J., Grosveld, F.G., Ren, B., Wendt, K.S., 2014. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. Proc Natl Acad Sci U S A 111, 996–1001. https://doi.org/10.1073/pnas.1317788111

ITQB NOVA, Avenida da República, 2780-157 Oeiras, Portugal | Tel +351 214 469 100 | itqb.unl.pt