UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA ANIMAL

Retinoic acid:
a key regulator of vertebrate embryonic development

Mestrado em Biologia Evolutiva e do Desenvolvimento

André Dias

Dissertação orientada por:
Doutor Moisés Mallo
Professora Doutora Sólveig Thorsteinsdóttir

2015

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA ANIMAL

Retinoic acid:
a key regulator of vertebrate embryonic development

Mestrado em Biologia Evolutiva e do Desenvolvimento

André Dias

Dissertação orientada por:
Doutor Moisés Mallo
Professora Doutora Sólveig Thorsteinsdóttir

2015

Resumo

O ácido retinóico é uma molécula sinalizadora, sintetizada a partir da vitamina A, necessária para o correto desenvolvimento embrionário dos vertebrados, uma vez que regula a transcrição de genes essenciais durante vários processos da embriogénese. A relação entre a sua produção e degradação, respetivamente através das enzimas Raldh2 e Cyp26a1, leva a que a sua atividade sinalizadora ocorra de uma forma bastante específica quer em termos de dose, localização da sua atuação (tecido) ou quer em termos temporais. De facto, a alteração dos níveis de ácido retinóico durante o desenvolvimento embrionário pode ser letal ou dar origem a doenças congénitas (como por exemplo, *spina bífida*)[1–6].

Uma das etapas mais importantes na embriogénese dos vertebrados é a gastrulação. Durante esta, para que a formação dos três folhetos germinativos (que mais tarde irão dar origem a todos os tecidos e órgãos do organismo) possa ocorrer devidamente, é necessário um correto controlo da atividade sinalizadora mediada pelo ácido retinóico, que nesta fase do desenvolvimento é fundamentalmente fornecido por via materna (caso dos ratinhos), dado que alterações na sua degradação enzimática pelas Cyp26 podem resultar na morte do embrião. Tal ocorre uma vez que o ácido retinóico controla a expressão de Nodal (excesso de ácido retinóico provoca a indução ectópica de Nodal), proteína indispensável para que se dê o início da gastrulação através da formação da linha primitiva, decorrente da migração das células do epiblasto para a parte mais posterior do embrião[7]. Aparentemente a introdução de um transgene (*T-streakCre^{ERT}*) juntamente com o alelo repórter *ROSA26R-β-gal* em ratinhos mutantes para o gene *Gdf11*, produziu letalidade embrionária durante a gastrulação[8]. Após verificarmos que o efeito não era devido ao transgene por si próprio, mas sim possivelmente ao local onde este foi inserido no genoma de uma linha particular (#47), procurámos caracterizar molecularmente os embriões *Gdf11$^{-/-}$::T-streak-Cre^{ERT}#47$^{+/0}$* de forma a tentar encontrar a causa desta aparente letalidade. A alteração dos níveis de ácido retinóico observada nos embriões *Gdf11$^{-/-}$* durante a transição entre a formação de tecidos do tronco e da cauda, fez-nos considerar a possibilidade de durante a gastrulação ocorrer uma alteração semelhante desses níveis, que combinada com possíveis efeitos do transgene pudesse resultar na morte dos referidos embriões. No entanto, as experiências realizadas não só não permitiram a identificação de qualquer problema durante a gastrulação, decorrente da mutação no gene *Gdf11* e/ou do transgene *T-streak-Cre^{ERT}#47$^{+/0}$*, como demonstraram a correta formação da linha primitiva nos embriões *Gdf11$^{-/-}$::T-streak-Cre^{ERT}#47$^{+/0}$*. Estudos adicionais mostraram ainda viabilidade dos embriões *Gdf11$^{-/-}$::T-streak-Cre^{ERT}#47$^{+/0}$* a E10.5. Uma possível explicação para este facto prende-se com a estratégia utilizada na genotipagem do alelo mutante de *Gdf11*, pois usava iniciadores desenhados para amplificar parte da cassete de neomicina (introduzida para criar a mutação nesse gene[9]) que também está presente no alelo repórter *ROSA26-β-gal*, que formava parte das experiências originais que levaram à nossa hipótese inicial. Assim sendo, é bastante provável que os erros cometidos na determinação do genótipo dos progenitores dos embriões *Gdf11::ROSA26R-β-gal$^{+/0}$::T-streakCre^{ERT}#47$^{+/0}$*, tenham levado a utilizar ratinhos *Gdf11$^{+/+}$* e não *Gdf11$^{+/-}$* nos cruzamentos, criando a ilusão de uma ausência de embriões *Gdf11$^{-/-}$* nestas ninhadas. De qualquer forma, relativamente à hipótese da alteração dos níveis de ácido retinóico nos mutantes *Gdf11*, foi possível concluir

que aparentemente essa alteração não se verifica durante a gastrulação, uma vez que a expressão tanto de *Cyp26a1*, como de *Nodal* permaneceram sem alterações.

Após a gastrulação, nos vertebrados, dá-se primeiro a formação da cabeça (na parte mais anterior do embrião) e depois por um processo de extensão axial, no sentido anterior para posterior, é formado o pescoço, a seguir o tronco e por último a cauda. Apesar do desenvolvimento destas estruturas ser progressivo e de depender dos progenitores axiais, parece que a forma como estas diferentes partes do corpo são criadas é diferente[10]. Mutações nos genes *T*, *Cdx* e *Wnt3a*[11–17] suportam esta teoria, segundo a qual a posição dos membros superiores e inferiores delimita os referidos blocos estruturais do embrião (cabeça/pescoço, tronco e cauda). Como referido anteriormente, o mecanismo segundo o qual ocorre a transição tronco-cauda já foi demonstrado pelo nosso laboratório[10], mas a transição entre a formação da cabeça e do tronco ainda permanece por esclarecer. Curiosamente a mutação da enzima *Raldh2*, que leva à inexistência de ácido retinóico nos tecidos neurais e na mesoderme, leva à morte do embrião após o desenvolvimento ser interrompido ao nível dos membros anteriores[18,19]. Mas, se for administrado ácido retinóico até essa altura do desenvolvimento (~E8.25) o embrião é capaz de ultrapassar esse bloqueio e formar as seguintes estruturas (tronco e cauda)[20]. Esse facto fez-nos considerar a hipótese de que o ácido retinóico pode estar a controlar o mecanismo de transição entre a formação da cabeça e do tronco. Para identificar essa necessária mudança nos progenitores axiais, dependente do ácido retinóico, efetuámos uma análise transcriptómica comparativa a partir de ARN isolado das caudas de embriões "tipo selvagem" a E8.75/E9.0 (ou seja, onde a transição já foi efetuada e estão a ser criados os tecidos do tronco) e de embriões *Raldh2*$^{-/-}$ da mesma idade (nos quais esta transição encontra-se bloqueada). Após análise dos resultados presentes na RNA-seq, foi selecionado um grupo de genes candidatos com base na grandeza da expressão diferencial observada entre embriões tipo selvagem e *Raldh2*$^{-/-}$, e tendo em conta a sua significância e função biológica. Vários genes desse grupo (exemplo: *Wnt3a*, *Dkk1 e Cav1*) estão associados à sinalização Wnt, cuja atividade canónica (via β-catenina) parece estar diminuída na cauda dos embriões mutantes. Esta observação é bastante interessante, tendo em conta a comparação dos fenótipos dos embriões mutantes para *Wnt3* (que apresentam ausência total de mesoderme), com o fenótipo dos embriões *Wnt3a*$^{-/-}$ (em que o desenvolvimento apenas ocorre de forma normal até ao nível do membro anterior)[21], sugerindo que a atividade do ácido retinóico possa ser responsável por esta mudança na sinalização Wnt (de depender de *Wnt3* e passar a depender de *Wnt3a*). Sendo que esta hipótese contrasta com a observação de que a expressão de *Wnt3a* parece estar aumentada na cauda dos mutantes[22], é possível que a existência de regulação diferencial dos vários componentes da sinalização Wnt nos embriões *Raldh2*$^{-/-}$ possa resultar na incapacidade dos progenitores de responder apropriadamente a Wnt3a, ocorrendo dessa forma a inibição da sinalização canónica de Wnt. Por esse motivo, através de hibridação in situ procurámos observar a expressão de alguns desses genes envolvidos na sinalização Wnt e complementámos esses estudos com uma abordagem de sobre-expressão através da utilização de transgénicos (onde o gene avaliado foi associado ao promotor de *Cdx2* que expressa nos progenitores do eixo). Estudos similares foram também feitos com outros genes não associados à sinalização Wnt (*Mesp1* e *Fgf4*). No entanto as nossas experiências não permitiram determinar o mecanismo inerente à mudança nos progenitores, necessária para que o embrião termine de criar tecidos da cabeça e inicie a produção de tecidos do tronco.

Ainda sobre a forma como a atividade do ácido retinóico influência os progenitores axiais, através de hibridação *in situ*, confirmámos que na mesoderme pré-somítica, a área onde tanto *T* como *Sox2* são expressos é menor nos embriões *Raldh2*$^{-/-}$. Em experiências preliminares foi possível observar, através de imunofluorescência, a existência de células que expressam T e Sox2 (apesar de não ter sido possível contabilizá-las), indicando dessa forma a presença dos progenitores axiais nos embriões *Raldh2*$^{-/-}$. Por último, surpreendentemente ao que está publicado, verificámos que em alguns embriões *Raldh2*$^{-/-}$ é possível observar umas pequenas protuberâncias onde *Tbx5* (marcador da indução dos membros anteriores) encontra-se expresso mais tarde no desenvolvimento (do que aquilo que seria suposto) e em menor quantidade. Por isso, considerando o fenótipo decorrente da mutação de *Cyp26a1* (onde ocorre uma transformação homeótica da vertebra cervical C5 para C6 e da C7 para uma vertebra torácica)[23], é possível que o ácido retinóico seja crucial durante este período do desenvolvimento, definindo o tempo em que ocorre a transição entre a formação da cabeça (pescoço) e do tronco no embrião. Apesar desta hipótese necessitar de uma validação experimental, conectando a transição cabeça-tronco e a indução dos membros anteriores, todas as evidências apontam no sentido de que a atividade do ácido retinóico nos progenitores axiais é apenas necessária durante a transição cabeça-tronco (e não durante todo o processo de extensão axial) sendo depois restringida, pela Cyp26a1, a zonas mais anteriores do embrião[24].

Palavras-chave

Ácido Retinóico, Raldh2, Transição Cabeça Tronco, Progenitores Axiais e Gdf11

Abstract

Retinoic acid (RA) is a signalling molecule, derived from vitamin A, necessary for proper vertebrate embryonic development. It acts in a tissue, time and dose specific manner, shaping the embryo through the regulation of several master transcription factors. Alterations in RA levels during embryonic development are known to cause several problems, including embryonic lethality[2].

During gastrulation, the formation of the three germ layers requires balanced interaction between RA and Cyp26 molecules, which is crucial for Nodal expression[7]. Interestingly, in Mallo's lab, embryos carrying a transgene (*T-streakCre*^{ERT}), the cre reporter allele *ROSA26R-β-gal*, and a total inactivation of *Gdf11*, seemed to die during this stage[8]. Considering the interaction between *Gdf11* and *Cyp26a1*[25], we hypothesized that the transgene, together with *Gdf11* might have affected RA signalling. However our experiments failed to find any indication of patterning problems that could justify the early lethality that was initially observed. Additional analyses indicate that the original phenotype might have resulted from incorrect genotyping of the parent lines. Also, we could conclude that despite what happens during the trunk to tail transition in *Gdf11*$^{-/-}$, in these mutants RA signalling seems not to be affected during gastrulation.

Embryos lacking RA (*Raldh2*$^{-/-}$) become truncated at the forelimb level[19]. Since Raldh2 mutant embryos exposed to acute RA treatments at E8.25 acquire trunk and tail structures[20] we concluded that RA signalling controls the mechanism regulating the head to trunk transition. To understand this process we performed a comparative transcriptomic analysis between tails of *Raldh2*$^{-/-}$ and wild type embryos. So far, we could not elucidate the mechanism for this transition but our evidence suggests that it possibly involves Wnt signalling. Also our experiments, concomitantly with data regarding *Cyp26a1*[23], seem to indicate that RA activity in the axial progenitors is only necessary during this transition, thus setting the time at which it occurs.

Keywords

Retinoic Acid, Raldh2, Head to Trunk Transition, Axial progenitors and Gdf11

Table of contents

Supplementary Information

I – Cuffdiff RNA-seq analysis (EXCEL file)

II – Solutions (PDF file)

III – Manufacturer protocols (PDF file)

IV – Plasmid maps and sequences (PDF file)

Chapter I

INTRODUCTION

General Introduction

I - Retinoic Acid

Retinoic acid (RA) is a signalling molecule, derived from vitamin A (retinol), necessary for proper embryonic development. Once retinol is inside the cell it is converted by retinol dehydrogenases (Rdh enzymes) into retinaldehyde, from where RA is synthesized through retinaldehyde dehydrogenases (Raldh enzymes). Endogenously produced or exogenous RA then binds to specific retinoic acid receptors, which interact with retinoic acid responsive elements (RAREs) in the genome, in order to activate the transcription of target genes **(Fig.1)**. Regarding these target genes, RA is known to regulate several master transcription factors necessary for key processes during formation of the vertebrate body (e.g. gastrulation and axial elongation). RA is also degraded by cytochrome P450 enzymes (e.g. Cyp26a1), which limits RA activity in a tissue, time and dose specific manner[1–3].

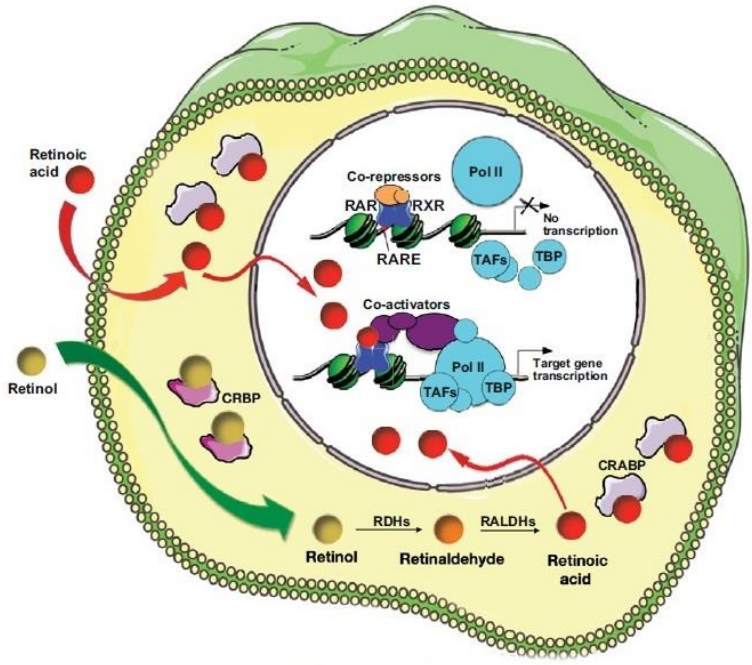


Fig. 1 – Summary of the retinoic acid signalling pathway, representing the activity of exogenous or intracellularly-produced RA as a transcriptional activator of gene expression. (From Rhinn and Dollé, 2012)

Alterations in RA signalling can produce a variety of problems during embryonic development, ranging from lethality to congenital spinal deformities in vertebrates (including spina bifida)[4–6]. Further below I will discuss several stages of vertebrate embryonic development to help better understanding how RA activity shapes the vertebrate embryo.

II - Vertebrate embryonic development – the mouse case

Vertebrates display a large diversity of body shapes and sizes. They are formed through multiple tightly regulated and interdependent morphogenetic events during embryonic development. Despite the gross architectural differences observed early in development across vertebrate species, the fundamental principles of vertebrate embryonic development are maintained throughout vertebrate phylogeny (e.g. gastrulation, somitogenesis and axial extension). However, the environment where embryonic development occurs represents one of the main differences among species; for instance, birds develop inside independent eggs, whereas mammals develop inside the progenitor uterus. I will now focus on mammalian embryonic development, using the mouse embryo as example[26,27].

II.a - From a fertilized egg to the gastrulating embryo

After fertilization, the egg undergoes cell divisions reaching the eight-cell stage, when apical-basal polarity is generated through a process known as compaction, forming a solid mass called morula[27,28]. Then the compacted embryo undergoes additional cleavages and cell fate decisions, eventually reaching the blastocyst stage[27,28]. At this stage it is possible to distinguish two different cell compartments, the trophectoderm (TE) and the inner cell mass (ICM). Cells from the TE will be necessary for implantation (which occurs around embryonic day (E) 4.5) and they will give rise to important extraembryonic tissues. The ICM contains pluripotent cells that undergo a second lineage separation, driven by Fgf signalling, to produce the visceral endoderm (positive for Gata4/6) and the epiblast (expressing *Nanog* and *Oct3/4*), which will give rise to the embryo proper[27–29]. Then, the epiblast changes morphologically to produce the so-called "egg cylinder" at around E5.5. At this stage the first signs of an AP axis are evident with the formation of the anterior visceral endoderm (AVE) (expressing Nodal and Wnt inhibitors). The AVE derives from the distal visceral endoderm (DVE) that migrates to the prospective anterior side of the embryo, shortly after being induced at the distal end of the embryo[27–31]. The molecular mechanism controlling the AP patterning of the embryo at this stage depends on interactions between β-catenin and Cripto. Other members of the Wnt and Nodal pathways, as well as Fgfs, are also involved in this important event[27,28,32]. Genetic experiments in the mouse and grafting experiments using other vertebrate model organisms showed that the AVE is involved in two main processes: the production of head structures later in the embryo and the induction/control of primitive streak (PS) formation in the opposite side of the egg cylinder, which will break the radial symmetry in the embryo and marks the onset of gastrulation[27,30].

II.b – The gastrulating embryo: formation of the primitive streak

Starting at E6.0 the mouse embryo undergoes a process called gastrulation. In this process, several cell movements rearrange the embryo to form the three germ layers (ectoderm, mesoderm and endoderm) that will eventually give rise to the different tissues and organs of the vertebrate body[27,30]. Maternally provided retinoic acid plays an important role at this stage by controlling *Nodal* expression in the epiblast. In particular, it has been shown that

when RA activity is not correctly buffered by Cyp26 enzymes, gastrulation fails[7]. Nodal is a protein belonging to the transforming growth factor β (Tgfβ) family responsible for sorting epiblast cells towards the posterior part of the embryo to generate the PS [33,34]. Wnt3-activated canonical Wnt pathway is also required for PS induction/maintenance and for the transcriptional activation of *Brachyury* (*T*) in the newly formed mesoderm[27,35–39]. PS induction requires expression of Wnt and Nodal inhibitors from the AVE to concentrate Nodal and Wnt signalling in the posterior epiblast. Accordingly, loss of these inhibitors (e.g. Cerl1 or Lefty1) resulted in the production of ectopic/enlarged PS[27]. Wnt/β-catenin signalling is also necessary for the maintenance of *Nodal* expression in the epiblast through a feedback loop involving *Nodal, Bmp4 and Wnt3*[40]. Epiblast cells in the PS will then undergo an epithelial to mesenchymal transition and ingress through the PS. These mesenchymal cells will give rise to the mesoderm and definitive endoderm[27,40]. This process also requires Fgf signalling (e.g. Fgf8) as its inactivation resulted in an accumulation of cells in the epiblast[41]. Epiblast cells located anterior to the newly formed PS are not affected by PS activity and therefore remain within the epiblast layer, eventually giving rise to the ectoderm[27].

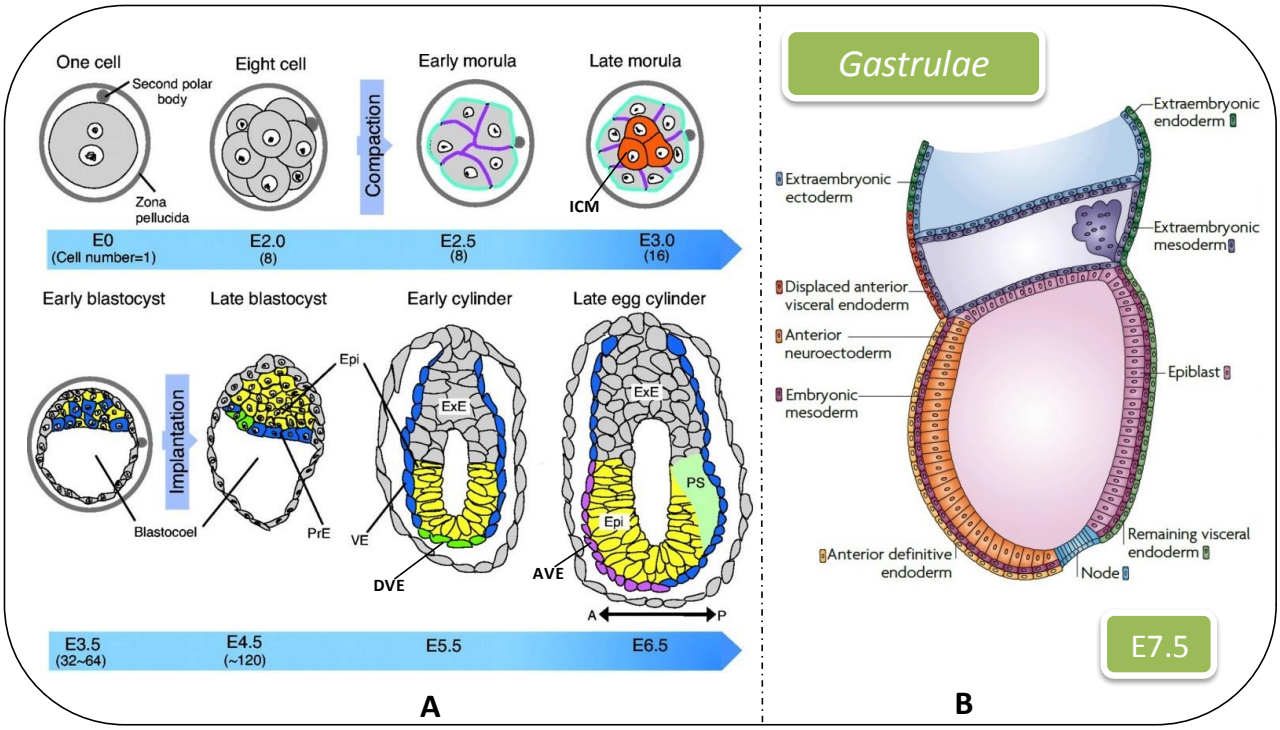


Fig. 2 – (A) Development of a fertilized egg into a gastrulating embryo; (B) Gastrulae: the three germ layer embryo. (Adapted from Takaoka et al, 2011, and Arnold et al, 2009)

II.c – The node and the left-right asymmetry

Using transplantation experiments, Spemann and Mangold discovered a group of cells in the amphibian embryo that have the capacity to induce the formation of a new vertebrate body axis. They coined the term organizer to describe this tissue, which is thought to be conserved across vertebrates. In the mouse it is referred to as the node[42]. This organizer is now considered as a secondary or later organizer, because the new axis it can induce does not

include formation head structures (only trunk and tail, since the induction of head structures depend on the AVE). At the end of gastrulation, the node can be observed at the most anterior end of the PS, which at this stage is fully extended. Fate mapping experiments indicated that cells ingressing through the node are fated to produce the notochord[43–45].

In addition to its role in gastrulation, the mouse node is also involved in the control of left-right asymmetry[45]. Disruption of the embryo's bilateral symmetry is mainly controlled by cells located in the node, which have motile cilia that rotate in a clockwise direction to create a leftward flow of extracellular fluid[43]. Experiments using mutant mice with immotile cilia showed a total absence of flow in the node that later in development results in a randomization of the organ *situs*. Also, an artificial reversal in the flow's direction was able to induce *situs inversus*. This flow seems then to direct some signalling activity from the node into the left side of the lateral plate mesoderm (LPM), once the reversal of the flow direction or its ablation results in changes of the organ *situs*[43,44]. The molecular components of this activity seem to vary among species (e.g. Shh in chicken[46]). However, Nodal signalling seems to be a key player in left-right asymmetry in most vertebrates. *Nodal* expression in the mouse can first be observed in the node during late PS and head-fold stages (~E7.75/E8.0). This expression is crucial for the creation of asymmetry in the embryo since blocking *Nodal* expression in the node inhibits later on *Nodal* expression on the left LPM, therefore creating left-right patterning defects[43,47]. A physical midline-barrier composed of the notochord and the floor plate, both derived from the node, has been proposed to maintain correct laterality of *Nodal* expression. The existence of a molecular midline-barrier, separating the expression of signals in the left-right LPM near the node has also been proposed based on studies of *Lefty1*$^{-/-}$ embryos. *Lefty* genes are expressed on the left LPM and are thought to cooperate with Nodal to orchestrate a left fate in the embryo[43,45]. However, the processes that lead to left-right asymmetric morphogenesis in the embryo are not fully understood. RA signalling might play a role in this process since different RA levels seem to alter *Nodal* expression in the LPM[48,49], thus creating several left-right patterning defects. Left-right alterations can also be observed in embryos lacking the RA synthesizing enzyme Raldh2, represented by asymmetric somite formation[50].

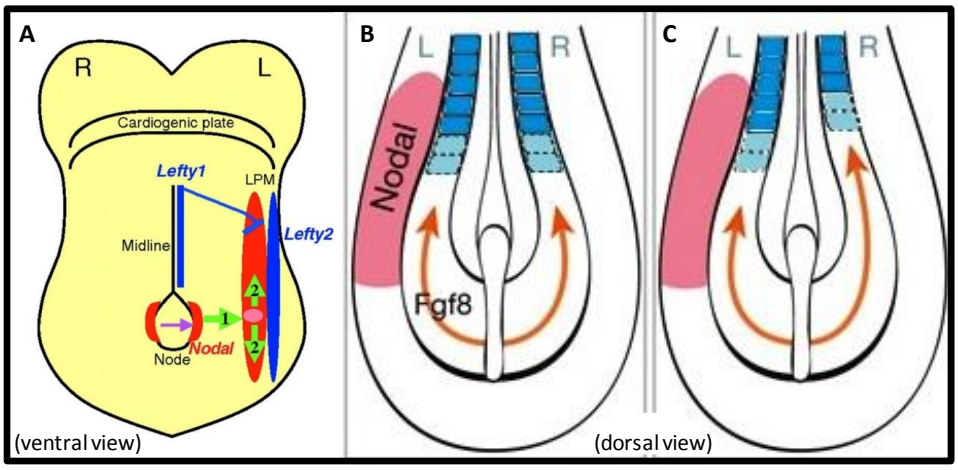


Fig. 3 – (A) Main molecular signallings responsible for the left-right asymmetry in an E8.0 embryo. **(B, C) Segmentation in wild-type embryos** (B) and in **Raldh2 mutant embryos,** which display asymmetric somite formation (C). (Adapted from Shiratori et al, 2006, and Vilhais-Neto et al, 2010)

II.d – From head to tail – axial elongation

After gastrulation, embryo growth is progressive. The first structure to form is the head at the anterior embryonic end. The axis then extends progressively in an anterior to posterior direction, first to produce the neck, then trunk structures and finally the tail[10,51]. Tightly associated with these processes is the continuous production of mesoderm, which seems to depend to a large extent on Wnt signalling. Loss of Wnt3 blocks mesoderm production, which can be evidenced by the absence of *T* expression[35,37,38]. This phenotype is similar to that of *β-catenin* mutants, suggesting that Wnt3 signals through the canonical pathway[39,52]. Later in development *Wnt3* expression begins to sag and, concomitantly, *Wnt3a* expression becomes activated at the posterior embryonic end. This change in Wnt ligand expression coincides with a switch in the Wnt molecule driving mesodermal production[53]. Accordingly, a null mutation in *Wnt3a* leads to strong axial truncation caudal to the forelimb level derived from the failure to produce mesoderm[17]. At this point, *T* expression, which initially depended on *Wnt3*, now requires *Wnt3a* signalling to be maintained. Interestingly, however, observations in *Lef1*$^{-/-}$*::Tcf1*$^{-/-}$ embryos suggest that this transcription complex downstream of Wnt3a is required for maintenance but not initiation of *T* expression. This seems to indicate that *Wnt3* and *Wnt3a* use a different set of effector complexes to regulate T expression[17,21,54,55]. Fgf signalling also plays a role in the regulation of *T* expression because *Fgf4* and *Fgf8* double mutants display reduced *T* expression in the axial stem zone[56].

Genetic experiments removing *T* and/or *Cdx* genes indicated that head structures are formed through a different process than trunk or tail. In particular, those experiments revealed that despite truncations in the main body axis these embryos were still able to produce head mesoderm and the first somites that originate cervical vertebra[11–16]. This important transition between head and trunk formation occurs at the level of the forelimb, which interestingly matches with the stage when *Raldh2*$^{-/-}$ stop developing (~E8.25)[18].

The progressive production of new tissues at the caudal end of the embryo relies on a pool of cells, known as the long term axial progenitors, located at the caudal tip of the embryo. These include the bipotent neural-mesodermal progenitors (NMPs) that give rise to the neural tube and the paraxial mesoderm[57,58]. Early in development, around E8.25/E8.5, NMPs are located at the node-streak border (NSB) and in the epiblast, between the node and the anterior PS. Later in development, after the mouse embryo underwent axial turning (~E9.5), NMPs are reallocated into the tailbud, to a region known as the chordoneural hinge (CNH)[58,59] – as roughly exemplified in **Fig.4**. Recent studies indicated that these cells co-express T and Sox2 and that they are able to self-renew within the embryo[58]. Also, they showed that *Tbx6* is required to drive the NMPs into a mesodermal fate, through a down-regulation of *Sox2*[60]. Accordingly, in the absence of *Tbx6*, embryos produce more neural tissue at the expense of paraxial mesoderm[61]. Other molecules, including Cdx proteins and signalling pathways activated by Wnt and Fgf ligands are also involved in the control of axial progenitors activity, although how is not fully understood[62–66].

Again RA plays a key role at different stages of axial elongation. A variety of experimental evidence indicates that RA is required for the tight balance between maintenance and differentiation of NMPs. RA activity is somehow required for the transition

from head to trunk development as revealed by the truncated phenotype of *Raldh2*$^{-/-}$ embryos[19]. Later in development, however RA signalling has to be kept on track (e.g. relocation of the NMPs into the CNH). This can be illustrated by genetic inactivation of Gdf11 signalling[9,25] or of *Cyp26a1* [23]. This last mutation, results in axial truncation at the level of the trunk to tail transition (TTT), due to excessive levels of RA, since ablation of *Rarg* rescued the caudal truncation in *Cyp26a1*$^{-/-}$ embryos[67]. Also, treatments with high RA doses during axial extension cause similar truncation phenotypes[68].

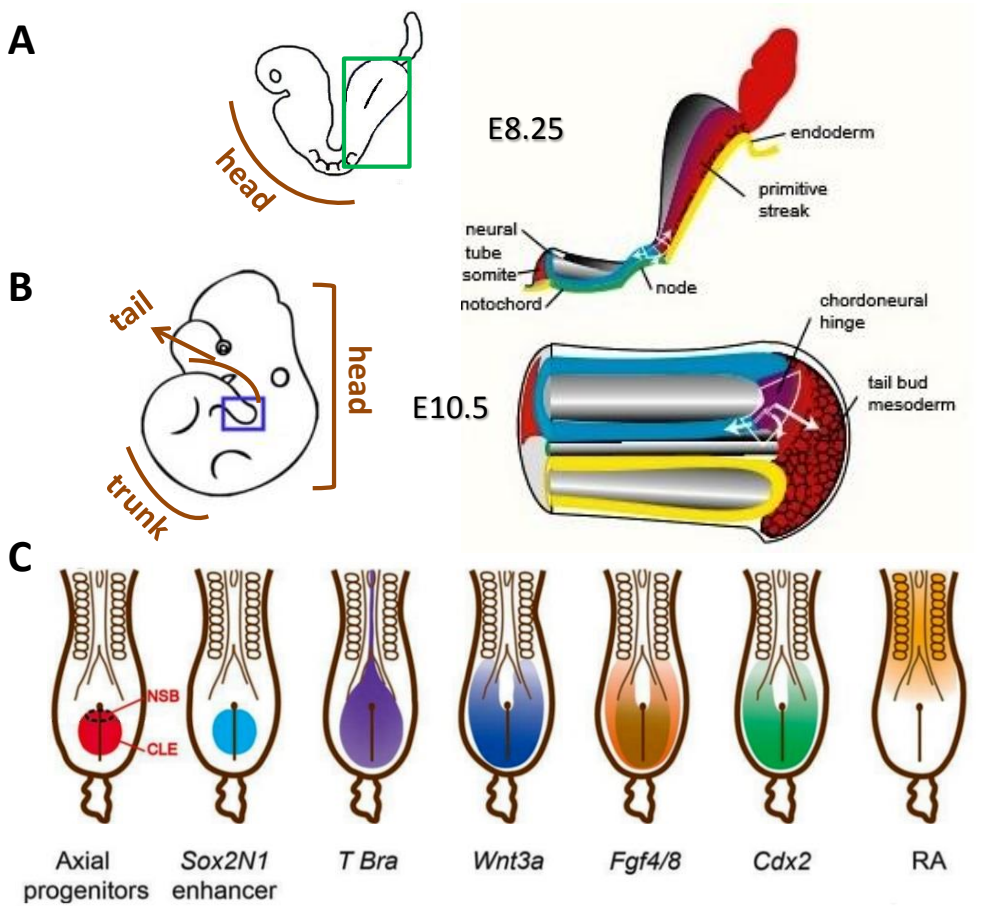


Fig. 4 – (A, B) Highlighted tailbud of an E8.25 (A) **and E10.5** (B) **embryo**, showing the location of NM progenitors in the NSB/epiblast (A) and later in development in the CNH (B). **(C) Expression domain of key molecular signals related with axial progenitors in an E8.5 embryo.** (Adapted from Cambray et al, 2002, and Neijts et al, 2013)

II.e – From progenitors to body structures

During axial extension, the body forms all the primordiae of the different tissues and organs of the body. These primordiae include pairs of symmetrical segments of mesodermal tissue at both sides of the neural tube, the somites, which will give rise to the axial skeleton, the dermis and skeletal muscles of the body and limbs. Somites are produced from head to tail according to a tightly rhythmic segmentation that takes place at the anterior border of the

presomitic mesoderm[51,69]. All vertebrates undergo somitogenesis but the number of formed somites and the time necessary for the formation of new segments varies across species. The periodic formation of somites relies on the existence of a molecular oscillator coined as the "segmentation clock"[69]. These molecular oscillations can be observed by cyclic expression of several genes in the PSM. Although the specific identity of these genes seems to vary across vertebrate species (e.g. *hairy1* in chicken[70]), most of them are members of the Notch signalling pathway or its target genes (e.g. *Mesp2*), whose requirement for the formation of inter-somitic boundaries seems to be conserved among vertebrates[69,71]. Apart from the "segmentation clock", somitogenesis also relays on another system, the "wavefront"[72] **(Fig.5)**. It consists of complex signalling gradients responsible for translation of signals from the molecular oscillator into patterning information, necessary for the formation of new somites[51,69,72]. This molecular control of segmentation is thought to derive from the convergence of two opposite functional gradients provided by Fgf and Wnt signalling (posterior to anterior inhibitory gradient) and RA signalling (anterior to posterior activating gradient), thus generating the determination front within the anterior PSM, where new intersomitic borders are created[69,71].

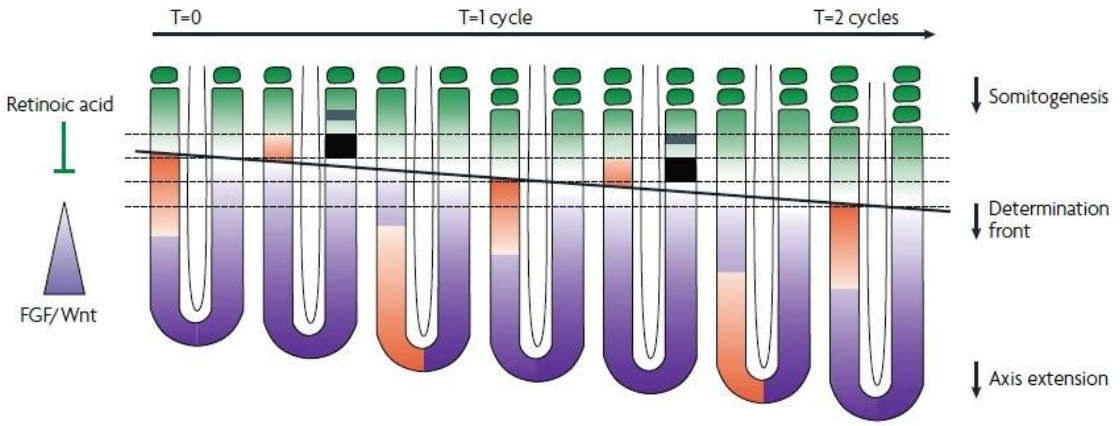


Fig. 5 – A somitogenesis model integrating the segmentation clock and the determination front. Opposing gradients, Fgf-Wnt (purple) and RA (green) sets the determination front (black line), in the PSM, during axial elongation. The wave of cyclic gene expression (orange) is represented on the left side of the embryo, whereas on the right side is represented the acquisition of a future segmental domain by *Mesp2* expression (in black). (From Dequéant and Pourquie, 2008)

As new segments are being produced, somites located at more anterior positions of the axis start differentiating. This differentiation incorporates positional information responsible for producing specific structures at the different axial levels in order to create a properly organized body[51,69]. In this process, AP patterning of the axial skeleton is mostly controlled by *Hox* genes[69,73] **(Fig.6)**. Several *Hox* clusters/genes have retinoic acid responsive elements (RAREs) nearby[18,74], which could explain the requirement of RA signalling for their proper expression and the many homeotic transformations resulting upon treatment with high doses of RA at different times of development[68,75].

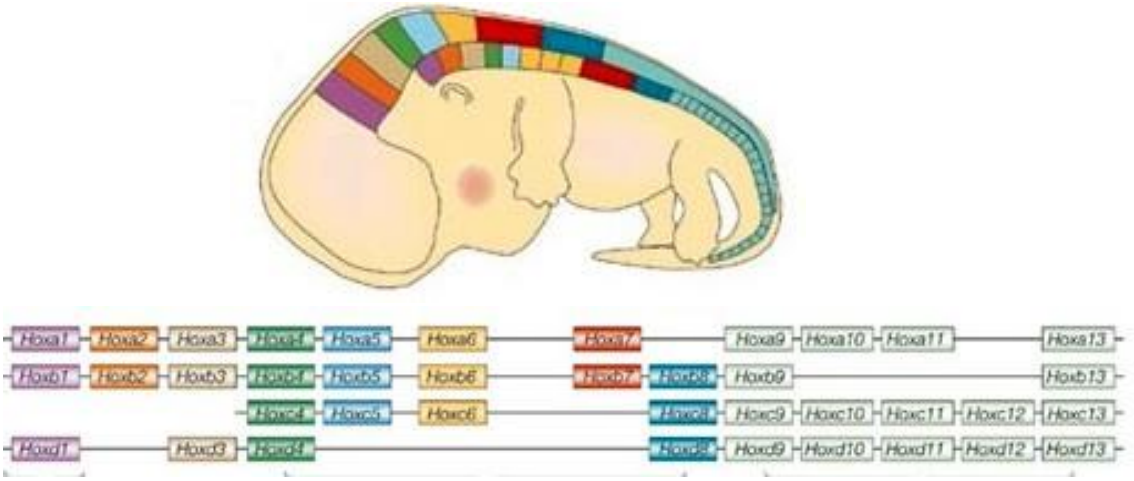


Fig. 6 – Hox gene expression and genomic organization in the mouse embryo; the paralog groups within the four clusters are color-coded according to their anterior-most expression domain in the mouse embryo. (Adapted from Pearson et al, 2005)

In addition to the formation of the body structures described above, RA is also important in organ formation. Heart and kidneys are examples of organs where RA is crucial for morphogenesis[3]. RA signalling is also involved in the induction and/or AP patterning of non-axial structures, like the forelimbs, where it is necessary for induction and initiation of the pre-limb bud[17,76,77] **(Fig.7)**.

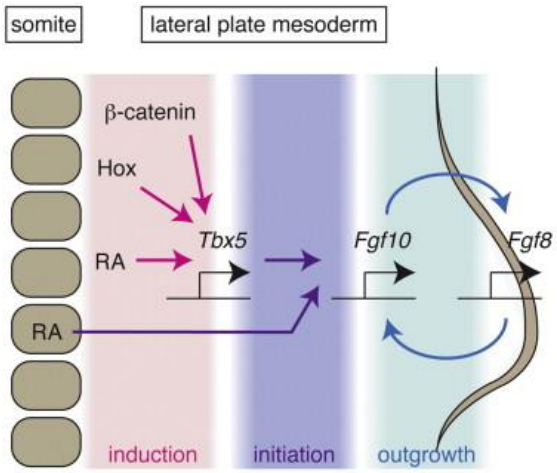


Fig. 7 – Model for forelimb bud induction, initiation and outgrowth. (From Logan et al, 2015)

In summary, RA is a key regulator of vertebrate embryonic development, required for a variety of process at different developmental stages. Importantly, its activity is often necessary during a specific time-window and in a dose and tissue dependent manner.

General Aims

The general aim of this MSc thesis was to contribute to a better understanding of how RA regulates vertebrate embryonic development. Two different projects were addressed during the period of this thesis and they are described here in *Chapter II and III*. Specific aims for each project of this thesis are specified in the referred chapters.

Chapter II

Retinoic acid and the mechanism involving the head to trunk transition in the axial progenitor cells

Retinoic acid activity during axial elongation

Studies using RA-responsive transgenic mice concluded that the Raldh2 enzyme is the main responsible for RA synthesis in neural and mesodermal tissues[19]. The genetic disruption of the Raldh2 enzyme leads to several abnormalities during development, after gastrulation, resulting in embryonic lethality at midgestation. This lethality could be rescued with RA treatments administered maternally at E8.25. *Raldh2* mutant embryos **(Fig 8)** display a development block starting around E8.25, failing to form limb buds and they do not undergo axial rotation (normally occurring at E8.25-E8.5). These embryos have an open neural tube, small otocysts, only the one branchial arch and the heart has only a single and expanded cavity[18–20]. In the paraxial mesoderm these mutants form only 10 to 12 somites, which are smaller and more densely packed than in normal embryos, which results in shortened AP axis. Interestingly, however, analysis of rescued embryos using RA treatment indicated that RA is only essential for correct somite formation until the 6th somite[18–20,78].

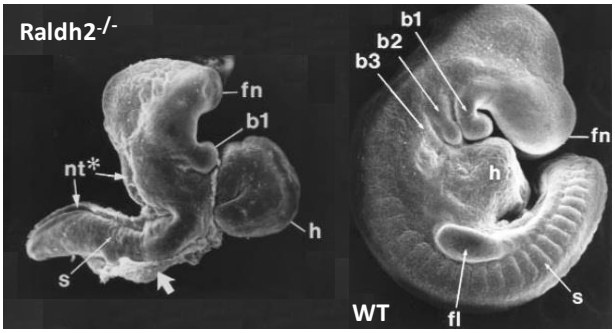


Fig. 8 – E9.5 *Raldh2* mutant versus wild type embryo (h- heart, b-branchial arch, s-somites, fn-frontonasal region, fl-forelimb bud, nt-neural tube). (Adapted from Niederreither et al, 1999)

As referred in *Chapter I, Raldh2*$^{-/-}$ show asymmetric development in the paraxial mesoderm around E8.25, during formation of the more anterior somites. This asymmetry is inverted in embryos with *situs inversus*. It is possible that this results from a requirement of RA to synchronize development in the paraxial mesoderm while maintaining asymmetric development in the lateral mesoderm[50,79–81]. Some of these patterning effects of RA have been shown to derive from interactions with Fgf8. In particular, it was found that RA antagonizes Fgf8 in the node ectoderm but not in the node mesoderm, where its expression is not uniform. In general RA sets the boundaries of *Fgf8* expression domain in the epiblast and the heart [18,50,74,82]. On the basis of these observations and considering the extended *Fgf8* expression domain in the presomitic mesoderm (PSM) of chicken embryos lacking RA signalling, it has been postulated that the asymmetry in the somites of *Raldh2* mutants derive from asymmetric *Fgf8* expression in the PSM[2,83,84]. The reciprocal interactions between RA and Fgf8 are also considered to be required for somitogenesis and axial extension by creating opposing gradients of activity in the PSM [6,22,63,69,84,85]. However, the relevance of continuous RA signalling

for somitogenesis and axial extension is a question of debate because *Raldh2*$^{-/-}$ can be rescued upon RA treatments at E8.25 and keep extending their axis and forming somites in the absence of RA. The fact is that all available evidence suggests that RA is fundamental at a specific time point during vertebrate embryonic axial extension, coincident with the stage when the forelimb buds are formed.

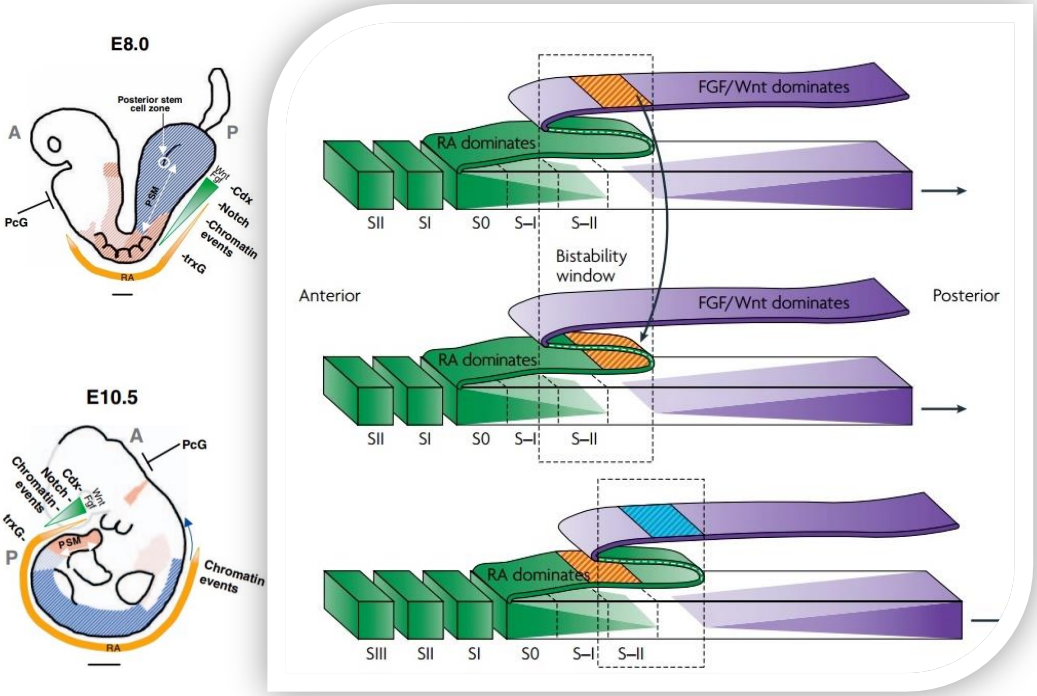


Fig. 9 – Model for Segment Determination (opposing RA and Fgf8 gradients during somitogenesis and axial entension). Adapted from Pourquie (2008) and Deschamps (2005)

Raldh2 mutant embryos are unable to induce forelimb buds as estimated by both the absence of a physical bud and of *Tbx5* expression, possibly due to an extension of the Fgf8 domains and/or to a direct role of RA production in somites in the regulation of Tbx5 in the pre-bud[18] . It has also been described that a strong and specific dose of RA (maternally administered) is necessary to induce forelimb outgrowth in these mutants[76,78]. Further analyses by Niederreither (2008) suggested that RA is essential for the induction and AP patterning of the pre-bud, but after this stage it must be removed for limb growth to continue normally. Interestingly, in rescued *Raldh2*$^{-/-}$ embryo hindlimb outgrowth was normal, indicating different RA requirements for forelimb and hindlimb development[76,78,86].

RA coordinates the Head to Trunk Transition?

As referred in *Chapter I*, vertebrate embryonic axial elongation relies on a pool of cells in the caudal tip of the embryo, characterized as long term axial progenitors[58]. In the last years, Mallo's lab have been studying the way these cells control axial growth and are responsible for the existent diversity of body shapes among vertebrates. Arnon Jurberg showed that Gdf11 signalling plays an essential role in the TTT[10]. This activity is in part mediated by the activation of *Isl1* in the NMPs, which will be responsible for the terminal differentiation of progenitors of the lateral mesoderm to produce the hindlimb[10]. More recently, Rita Aires during her PhD work *(unpublished)* observed that the tailbud of *Gdf11*$^{-/-}$ embryos contained some *Oct4* expressing cells (much later than normal). This persistent *Oct4* expression seems to be responsible for the increased body length of *Gdf11*$^{-/-}$ embryos since when *Oct4* was transgenically overexpressed in the tailbud, under the regulation of a *Cdx2* promoter, embryos had longer trunks and severe trunk to tail transition defects[87].

The phenotypic effects caused by inactivation of a variety of genes (e.g. *T*, *Cdx1/2/4* and *Wnt3a*[11–17]) indicate that during vertebrate body formation, the mechanisms supporting head, and trunk development are also fundamentally different, indicating the existence of mechanisms specifically regulating the head to trunk transition (HTT). Interestingly the stage at which RA signalling seems to be fundamental during embryonic body axis formation, matches with the time when the production of head tissues stops and the formation of trunk tissues starts. So the specific aim of this project was to address the hypothesis of possible activated RA targets, in the node streak border, be responsible for a switch in the long term axial progenitors necessary for the HTT in the vertebrate embryo.

Material and methods

[§]Solutions detailed information is present in **Supplementary Information II**

*All manufacturer protocols are present in **Supplementary Information III**

[#]Plasmid maps are shown in **Supplementary Information IV**

Mice/Embryos

All experiments and procedures conducted on mice followed the Portuguese (Portaria 1005/92) and European (Directive 2010/63/EU) legislations, concerning housing, husbandry, and welfare. These animals were kept in 12h dark/light cycle, some maintained in C57BL/6J background at the Rodent Facility and others (necessary for producing transgenic embryos) in FVB/N strain at the Pathogen-free Animal Facility, also at the IGC.

The transgenic embryos (*Cdx2.Cav1T1* and *Cdx2.Cav1T2*) were generated using standard transgenic procedures (e.g. Nagy et al 2003), through pronuclear microinjection by the Transgenic Unit at the IGC.

The TOPGAL transgenic mice, *Tg(TCF/Lef1-lacZ)34Efu/J*, described in Gupta and Fuchs (1999) were purchased from Jackson Laboratory.

RA mutant mice were created by producing a *Raldh2* allele unable to produce the protein product, by the introduction of stop codons in the second exon using the CRISPR-Cas9 technique. The mutation was confirmed by sequencing. $Raldh2_{Ext}$ designed oligos were used to amplify part of the *Raldh2* coding sequence from genomic DNA obtained from the tails of the transgenic progeny. The band containing the mutation was isolated in a 1% agarose gel in TAE using the QIAEX II Gel Extraction Kit*. Molecular cloning procedures were performed as described further below. In this case, the amplified fragment was digested with XhoI and XbaI, and inserted into the XhoI and SalI sites of the pKXM plasmid[#]. Plasmids containing the insert were sequenced to confirm the introduction of the mutation. The lethality of the mutation was also phenotypically confirmed since no $Raldh2^{-/-}$ pups were found upon crosses between $Raldh2^{+/-}$ mice, and the phenotype of $Raldh2^{-/-}$ embryos matched with what is described in the bibliography.

Matings were done late in the afternoon and plugs were checked in the morning of the next day (corresponding to E0.5). Embryos were collected by caesarean section on cold PBS[§] and fixed in 4%PFA[§], at 4°C overnight. They were then washed in PBT[§] and dehydrated in graded methanol series (25%, 50%, 75% made in PBT), washed in 100% methanol and stored at -20°C in methanol.

PCR conditions and genotyping

During the embryo harvesting process, the yolk sac was isolated and incubated overnight (ON) at 50^{0}C, in yolk sac lysis buffer[§] containing proteinase K. Genomic DNA was isolated from mouse tail biopsies and incubated under the same conditions, but in PBND buffer[§] containing proteinase K. After the ON incubation, proteinase K was inactivated at 95^{0}C for 15 min. Primers for several different PCR reactions and genotyping conditions are described in **Tables 1 and 2**. Exceptions are referred to in the corresponding sections.

Primers for	Forward	Reverse
Raldh2 $_{Mut}$	GTTTTCTGATCTCCCAGATCTC	TCTCCCACTGAATTCTATCAC
Raldh2 $_{WT}$		AACACTCTCCCACTCTCTGAG
Raldh2 $_{Ext}$		CATCTTCTAAGCAATACACAC
Raldh2 $_{Clon}$	GACTCGAGTTTTCTGATCTCCCAGATCTC	GATCTAGATCTTCTAAGCAATACACAC
Cav1.T1 $_{ORF}$	CTGTCGACCTCCTCAGAGCCTGCAGCCAG	GAGCGGCCGCGTCCCTCATATCTCTTTCTGCG
Cav1.T2 $_{ORF}$	CTGTCGACTGTTCCCATCCGGGAACAG	GAGCGGCCGCGTCCCTCATATCTCTTTCTGCG
Cav1.T1 $_{RT}$	CTGTCGACCTCCTCAGAGCCTGCAGCCAG	TTCTGGTTCTGCAATCACATC
Cav1.T2 $_{RT}$	ACAGCCAGGCTGACTCTTGAC	TTCTGGTTCTGCAATCACATC
Mesp1 $_{ORF}$	CCGTCGACGGATAAAGCTACAGCGGACCC	GAGCGGCCGCCAAAGGAAAAGTGTCTGTGC
Mesp1 $_{Bridge}$	CAGTCCCTCATCTCCGCTCTTCAGCAGCGACATGCTG	GTCGCTGCTGAAGAGCGGAGATGAGGGACTGGGCTCC
Mesp1 $_{RT}$	CGCAGAAACAGCATCCCAGG	TGTCCCCTCCACTCTTCAGGC
Eno2 $_{RT}$	CAAGCTGGCCCAGGAGAATGG	CTGGTTGTACTTCGCCAGACG
Fgf4 $_{RT}$	CCGGTGCAGCGAGGCGTGGTG	GTACGCGTAGGATTCGTAGGCG
Dkk1 $_{ORF}$	TGCGTCCTTCGGAGATGATGGTTG	CTGTCGGTTTAGTGTCTCTGGCAG
Actin $_{RT}$	ATGAAGATCCTGACCGAGCG	TACTTGCGCTCAGGAGGAGC
TOPGAL	CGTGGCCTGATTCATTCC	CGTGGCCTGATTCATTCC

Table 1 –Primers used for polymerase chain reaction
(normal PCR, RT-PCR and RT-qPCR).

PCR reaction contained	Quantity
Template (DNA)	~ 1µL
Primer Forward	0,25µL
Primer Reverse	0,25µL
dNTPs 25mM	0,2µL (0,2mM)
MgCl2 25mM	2,5µL (2,5mM)
Taq Buffer 10x	2,5µL (1x)
Taq polymerase (NZYTECH or Fermentas) 5und/µL	0,2µL (1und)
H_20 mili-Q	up to 25µL

<table>
<tr><th colspan="3">Oligos</th><th colspan="3">PCR conditions</th></tr>
<tr><td>Raldh2 Mut</td><td colspan="2">Raldh2 WT</td><td rowspan="10">95°C for 5 min</td><td>29 cycles (95°C for 45 sec, 60°C for 45 min and 72°C for 1 min)</td><td>72°C for 5 min</td></tr>
<tr><td>Raldh2 Ext</td><td>TopGal</td><td>Cav1.T2 RT</td><td>34 cycles (95°C for 45 sec, 62°C for 1 min and 72°C for 2 min)</td><td rowspan="5">72°C for 10 min</td></tr>
<tr><td rowspan="4">Cav1.T2 ORF</td><td colspan="2">Mesp1 ORF (Forward)</td><td rowspan="4">40 cycles (95°C for 45 sec, 62°C for 1 min and 72°C for 2 min)</td></tr>
<tr><td colspan="2">+</td></tr>
<tr><td colspan="2">Mesp1 Bridge</td></tr>
<tr><td colspan="2">(Reverse)</td></tr>
<tr><td>Raldh2 Clon</td><td>Cav1.T1 RT</td><td>Dkk1 ORF</td><td>34 cycles (95°C for 45 sec, 60°C for 1 min and 72°C for 2 min)</td></tr>
<tr><td rowspan="4">Cav1.T1 ORF</td><td colspan="2">Mesp1 ORF (Reverse)</td><td rowspan="4">40 cycles (95°C for 45 sec, 60°C for 1 min and 72°C for 2 min)</td></tr>
<tr><td colspan="2">+</td></tr>
<tr><td colspan="2">Mesp1 Bridge</td></tr>
<tr><td colspan="2">(Forward)</td></tr>
<tr><td colspan="3">Extension PCR (no oligos used#)</td><td>4 cycles (95°C for 45 sec, 60°C for 45sec and 72°C for 1,5 min)</td><td>72°C for 5 min</td></tr>
<tr><td colspan="3">Mesp1 ORF (after extension PCR#)</td><td>37 cycles (95°C for 45 sec, 60°C for 45sec and 72°C for 2 min)</td><td>72°C for 10 min</td></tr>
</table>

used for Mesp1 cloning

Table 2 – PCR conditions

RT-PCR

To obtain the desired cDNAs from isolated RNA (described below), a Reverse Transcriptase reaction was performed using the NZYTECH RT Kit* according to manufacturer's instructions. In these reactions, we used random hexamers for priming and the NZY Ribonuclease Inhibitor was substituted by nuclease-free water. At the end of the incubation, the cDNA was stored at -20°C.

PCR reactions were then performed using about 4µL of cDNA under the conditions specified in **Table 3**.

<table>
<tr><th colspan="5">Oligos</th><th colspan="3">RT-PCR conditions</th></tr>
<tr><td>Cav1.T1 RT</td><td>Cav1.T2 RT</td><td>Eno2 RT</td><td>Mesp1 RT</td><td>Actin RT</td><td rowspan="2">95°C for 5 min</td><td>40 cycles (95°C for 45 sec, 60°C for 1 min and 72°C for 2 min)</td><td rowspan="2">72°C for 10 min</td></tr>
<tr><td colspan="5">Fgf4 RT</td><td>40 cycles (95°C for 45 sec, 62°C for 1 min and 72°C for 2 min)</td></tr>
</table>

Table 3 – Specific RT-PCR conditions

qPCR

SYBR Green quantitative PCR analysis* was performed using cDNA obtained from RNA isolated from tails of Raldh2 embryos, according to the manufacturer's protocol and under the conditions described in **Table 4**. The results were analysed by Bio-Rad CFX Manager software

Plate Read	95^0C	5 min	
	95^0C	10 sec	39 cycles
	60^0C	30 sec	
	72^0C	30 sec	
Melting Curve	65^0C	5 sec	
	95^0C	increment $0,5^0$C	

Table 4 – qPCR conditions

Agarose gel electrophoresis

Agarose was dissolved in 1X TAE[§], usually at the concentration of 0.8%, 1% or 2%. Ethidium bromide or GelRed (1:39 in H_2O) was added ~1:4 in order to visualize the DNA when UV light was applied. 6x Gel loading dye was added to each sample (1x final concentration). An electric current of about 120V was applied to the gel immersed in 1X TAE. The QIAEX II Gel Extraction Kit* was normally used to extract the DNA from the agarose gels (eluted in TE[§]).

Phenol-Chloroform extraction and standard digestions

In several situations described below, DNA was purified by phenol-chloroform extraction. For this, TE buffer was used to make a final volume of 100µL and an equal volume of phenol-chloroform was added. The sample was mixed and centrifuged for 3 min at 14000 rpm. The DNA was recovered from the aqueous phase and precipitated with a 1:10 volume of 3M NaOAc pH 5.3 and 2.5 volumes of 100% ethanol for 30min on dry ice. The precipitated DNA was recovered by centrifugation at 14000rpm for 30min at 4ºC. The retrieved (air-dried) DNA pellet was resuspended in an appropriate volume of water or TE for further experiments (DNA concentrations were determined with a Nanodrop). Standard digestions using restriction enzymes were some of those applications. For that, to 5µL of DNA we normally added to 13 µL of H_2O, about 0,5µL of Enzyme and 2µL of the 10X concentrated buffer. The resulting mixture was incubated at least during 1h 30 min at 37^0C.

RNA extraction from the tails of *Raldh2* embryos

To analyze the transcriptome of axial progenitor cells, we isolated tails from E8.75/E9.0 *Raldh2* embryos (resulting from *Ralhd2*$^{+/-}$ mouse intercrosses), which were stored immediately at -80^0C. Upon genotyping, tails were grouped according to their category in groups of about 8 tails, as shown in **Fig. 10**. RNA was then isolated from these selected tissues, including two replicate groups for each category (WT and MUT) using TRI Reagent (Trizol), under the

conditions specified in the TRI REAGENT SIGMA protocol*. 11µL of nuclease-free water was added to the RNA pellet, which was dissolved at 65^{0}C during 10 minutes and then put on ice; finally samples were stored at -80^{0}C.

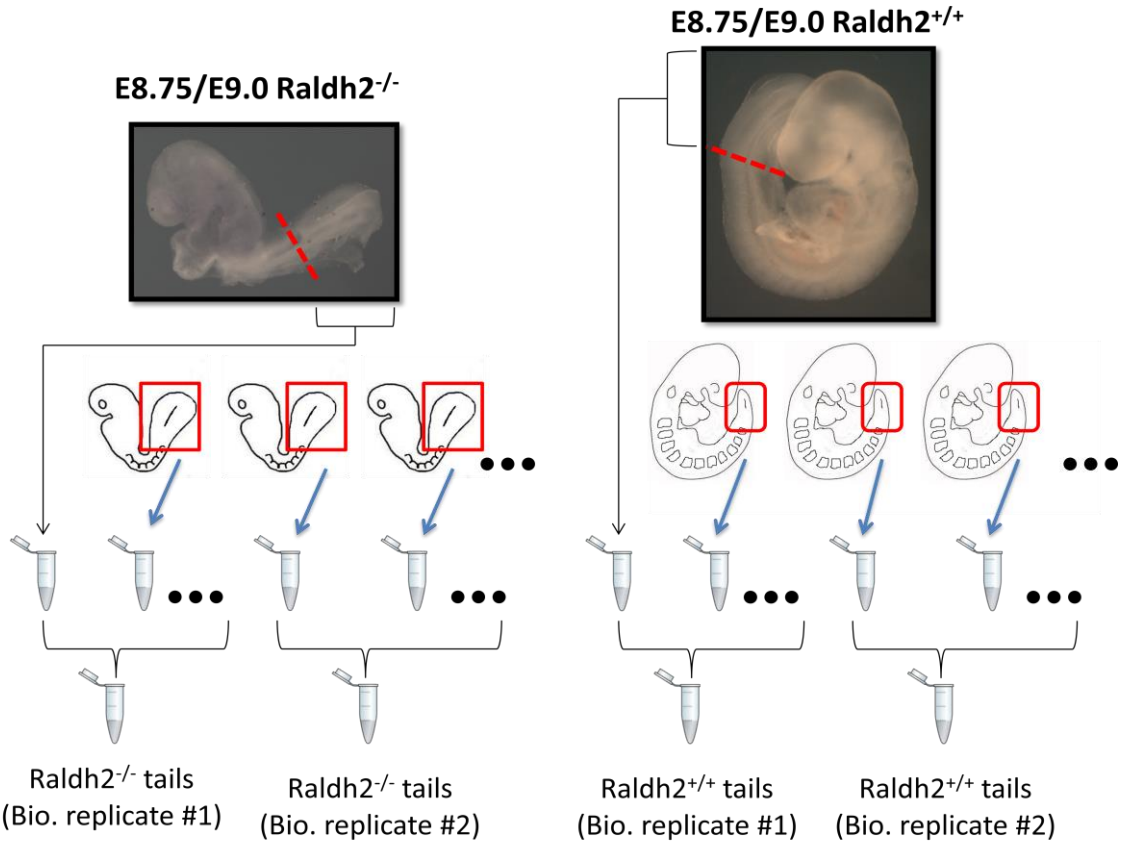


Fig. 10 – Process for Raldh2 isolation and storage in biological groups and category (Raldh2$^{+/+}$ or (Raldh2-/-)

RNA-seq

After the RNA was isolated from the selected group of tail tissues, a small amount was sent to the Gene Expression Unit (IGC), where the sample quality was assessed. Once the RNA integrity and concentration had the necessary levels, the four samples (RNAmut1, RNAmut2, RNAwt1 and RNAwt2) were sent to the EMBL in Germany where the RNA-seq was conducted. The RNA-seq results were analysed by the Bioinformatics Unit (IGC) using the Cuffdiff algorithm, which estimates expression at transcript-level resolution and controls for variability across replicates.

Molecular cloning

Specific DNA sequences, amplified by PCR (and purified using phenol-chloroform), and the chosen plasmids (vectors) were digested with the appropriate restriction enzymes before they were loaded in a 1% agarose gel in TAE and the bands separated by electrophoresis. The bands of interest were isolated and purified using the QIAEX II Gel Extraction Kit*. Then the

sample concentration was measured using a Nanodrop and the ligation reactions were performed with T4 DNA ligase using vector and insert in a 1-5 proportion. After 1h at room temperature the transformation was performed using DH5-α competent cells in a 1/10 reaction on ice followed by a heat shock (42°C) during 45 seconds and chilled on ice before being grown in LB medium at 37°C for 1 hour. Finally the bacteria were plated on solid LB medium with ampicillin (50 µg/mL) and incubated ON at 37°C. Single colonies were then picked and grown on LB medium with ampicillin at 37°C, during 3 hours with shaking. These cultures were used for screening purposes (see next section) and stored at 4^0C.

Molecular screenings and Plasmid DNA, Mini- and Midi-scale preparations

To perform a fast screen for positive colonies, 1µL of the above cultures was used for a PCR reaction, using the appropriate primers for each case. For colonies giving a positive PCR signal, 20µL of the culture was retrieved and added to 5mL of LB+ampicillin. Then, 1,5mL of that culture was spun down and the pellet resuspended in 100µl of TE with RNase (10µg/mL). Then, 300µl of TENS[§] and 150µl of 3M KOAc (pH5,2) were added to the sample. After mixing, the mixture was centrifuged (4min at 14000rpm) and the supernatant was transferred into a fresh tube containing 900µl of 100% EtOH. Then a spin was performed to pellet DNA and RNA, which were dissolved in 50µL TE in order to be used in screening digestion reactions. When higher purity DNA was required, plasmids were purified using commercial plasmid preparation kits: "NZYTECH MINIPREP" kit* was used when small amounts of plasmid were needed (for sequencing reactions) and "MN Plasmid DNA purification (NucleoBond Xtra Midi)" kit* was used when larger amounts of plasmid DNA were necessary (e.g. for later in vitro transcription experiments).

Sequencing reactions

To confirm the sequences of the cloned DNA products, cycle sequencing reactions were performed as described below **(Table 5)**.

Reagent	**Quantity**
Template (plasmid DNA mini prep)	~3µL (normally about 350ng)
Primer (T7 or T3)	1 µL (~5pmol)
Buffer	2µL
Terminator	2µL
H20 Mili-Q	Up to 10µL

96^0C	1 min	
96^0C	10 sec	25 cycles
50^0C	5 sec	
60^0C	4 min	
4^0C	-	

Table 5 – Standard cycle reaction conditions

The amplified DNA was then precipitated with 2µL of 3M NaOAc and 50µL of 95% EtOH (plus 10µL H_20) for 30 min at RT, centrifuged during 30min at 4ºC, the pellet was washed with 250µL of 70% ethanol and finally centrifuged at 14000 rpm for 15min at 4ºC. The supernatant was again removed and the pellet air-dried. The samples were then sent to the Genomics Unit at the IGC. The output sequences were analysed by the combined use of Finch TV software and BLAST (NCBI).

Microinjection construct

20 µg of the DNA construct was digested with the appropriate restriction enzymes to remove plasmid sequences and gel purified using the QIAquick Gel Extration Kit*. The DNA was eluted with 40µL of buffer EB and stored at -20ºC.

Probe generation by in vitro transcription

RNA probes for in situ hybridization were produced by transcription in vitro. First, a specific restriction enzyme was used to linearize 10µg of plasmid containing the relevant cDNA, purified by a phenol-chloroform extraction and resuspended in 20µL of water. Then, about 1µg of cDNA was used for RNA transcription **(Table 6)** for 3h at 37^0C. The RNA was recovered by NaOAc precipitation and its length confirmed in an agarose gel (normally 2%).

DNA	~1ug
T7/T3 polymerase	1µL
Buffer 10x	2µL
DIG label	1µL
Nuclease free water	Up to 20µL

Table 6 – RNA transcription reaction

Cav1 cloning, probe and transgenic construct

Cav1 has two transcript variants, *Cav1.T1* representing the longer one and *Cav1.T2* representing the shorter one. Cav1.T1 was amplified by PCR using cDNA produced by reverse transcription of RNA isolated from ES cells. The PCR reaction used Pfu polymerase, 5mM $MgCl_2$ and $Cav1.T1_{ORF}$ oligos. *Cav1.T2* was amplified (also with Pfu and with $Cav1.T2_{ORF}$ oligos) using the Cav1.T1 amplified sequence as template. They were inserted into the SalI and NotI sites of pBluescript II KS# using standard cloning procedures.

To create the *Cav1.T1* and *Cav1.T2* transgenic constructs under the *Cdx2* promotor, the cloned coding sequences were retrieved from the initial pKS plasmid with SalI and NotI and cloned into these sites of the pKS+Cdx2promotor+polyA plasmids#. The final specific constructs for microinjection were obtained through digestion with ClaI, according to the above-described methodologies.

For in vitro transcription, the *pKS.Cav1.T1* plasmid was linearized with SalI and the probe was synthesized with T7 RNA polymerase.

Dkk1 probe

The *Dkk1* coding sequence was amplified using cDNA obtained from RNA extracted from *Raldh2*$^{-/-}$ embryo tails. For this PCR, we used HotStart Taq* (QIAGEN) under the manufacturer conditions. The cloning was made with a PCRII-TOPO plasmid# using the TOPO TA cloning Kit*. Standard molecular cloning procedures were adopted and the plasmids transformed into DH5α competent cells. For in vitro transcription, the plasmid was linearized with SpeI and the probe was synthesized with T7 RNA polymerase.

Mesp1 probe

Once the entire coding sequence of *Mesp1* is split by just one intron we designed oligos to amplify the cDNA from genomic DNA, linking the two exons in vitro. For this we performed two separated PCRs: "A" using *Mesp1*$_{ORF}$ *(Forward) + Mesp1*$_{Bridge}$*(Reverse)* oligos and in "B" *Mesp1*$_{ORF}$ *(Reverse) + Mesp1*$_{Bridge}$, both using Pfu polymerase. Then, after the amplified DNA was gel purified with QIAEX Gel Extraction Kit, equimolar amounts were mixed and used as a template for PCR cycles without primers, also using Pfu polymerase. Finally, we added the *Mesp1*$_{ORF}$ *(forward and reverse)* oligos to the PCR reaction. After the required band was isolated from a 1% agarose gel, we followed the standard cloning techniques to introduce the *Mesp1* coding sequence into the SalI and NotI sites of pKS bluescript (BLAST was performed against *"Mus musculus mRNA for MespI, complete cds", NCBI accession: D83674)*. For in vitro transcription, the plasmid was linearized with SalI and the probe was synthesized with T7 RNA polymerase.

Other probes

All the other probes used in this work were available in the lab synthesized by in vitro transcription.

In situ hybridization

All the whole mount in situ hybridizations were performed using DIG-labeled antisense RNA probes as described in Kanzler et al 1998. On the first day, the embryos were rehydrated, and washed in PBT. They were then bleached in 6% H_2O_2 at RT for 1 h and washed in PBT. They were then treated with proteinase K for time lengths that depended on their developmental stage. After inactivating proteinase K with glycine§ and several washes in PBT, the embryos were post-fixed with PFA/glutaraldehyde§. Pre-hybridization solution§ was then added and incubated for 1 hour at 65°C, after which it was changed for hybridization solution containing the probe (3 to 6µL of probe per mL of pre-hybridization solution), and incubated overnight

between 65ºC and 70ºC. On day 2, after several washes using a post-hybridization solution§ and TBST§ the embryos were incubated in blocking solution (MABT/block/10% sheep serum§), first without antibody during 2,5h and then with the antibody against DIG (1:2000), overnight at 4ºC. On day three several washes were performed with MABT and on day four the embryos were incubated with a developing solution (NTMT§ plus NBT/BCIP), at RT, protected from the light. The reaction was stopped with PBT, the embryos fixed in 4% PFA and finally stored in PBT.

Whole mount embryo immunofluorescent staining

After rehydration from methanol to PBS, the embryos were washed 4x30 min in PBST§, incubated with 1M glycine in 0,1% PBST for 30 min to reduce unspecific binding and washed 3x in PBST to remove glycine residues. The embryos were then blocked in donkey serum blocking buffer§ overnight at 4^{0}C. On the next day, the embryos were incubated ON at 4^{0}C with new blocking serum containing (~1:250) anti-T *(Goat AF2085 from R&D)* and anti-Sox2 *(Rabbit ab92494 monoclonal from abcam)* primary antibodies. On the third day, after several washes in PBST, the embryos were incubated in new blocking serum containing the secondary antibodies (donkey anti goat rabbit and donkey anti rabbit, ~1:1000). The last day, after several washes of PBST, the embryos were incubated with DAPI in PBST (~1:5000) during 2,5 hours and finally through a process involving graded washes in methanol to methyl salicylate to clear the embryos, which were prepared in a blade for confocal microscopy.

Wnt reporter activity

To observe β-catenin signalling activity in *Raldh2* mutant embryos, we introduced the TOPGAL Wnt reporter mice into the *Raldh2*$^{+/-}$ background. *Raldh2*$^{+/-}$*::TOPGAL*$^{+/0}$ males were crossed with *Raldh2*$^{+/-}$*::TOPGAL*$^{+/0}$ females to obtain *Raldh2*$^{-/-}$*::TOPGAL*$^{+/0}$ embryos, which were fixed in Mirky's ON at 4ºC. After 3x10 min washes with 0,02%Tween-20/NP40 in PBS, the embryos were stained (protected from the light) at 37^{0}C with X-gal staining solution§ and finally post fixed ON, at 4ºC, with Mirky's§.

Looking for a change in the axial progenitors

Previous studies indicated that RA signalling is essential for the vertebrate embryo to undergo HTT. We thus decided to use *Raldh2* mutant mouse embryos, which lack neural and mesodermal RA activity, to search for the mechanisms controlling this transition. In the mouse embryo this transition occurs around E8.25/E8.5, roughly corresponding to the stage when the forelimb bud is induced. We therefore isolated tails from wild type E8.75/E9.0 embryos, which already started trunk development, and tails from *Raldh2*$^{-/-}$ littermates, which display a strong developmental delay, possibly resulting from a failure to undergo the HTT. We then analysed gene expression in these tissues by RNA-seq and compared their mRNA profiles **(Fig. 10**, **11** and **Supplementary Information I)**.

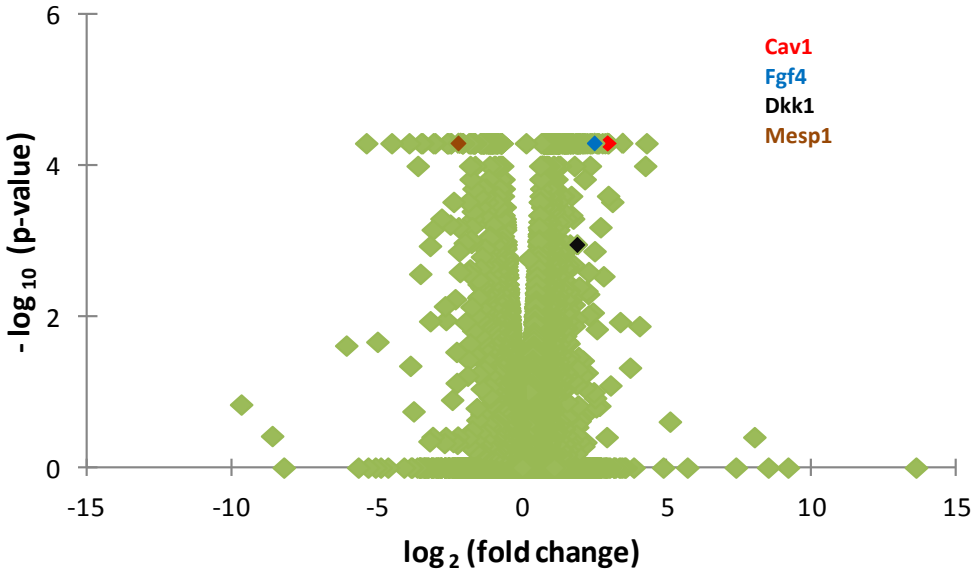
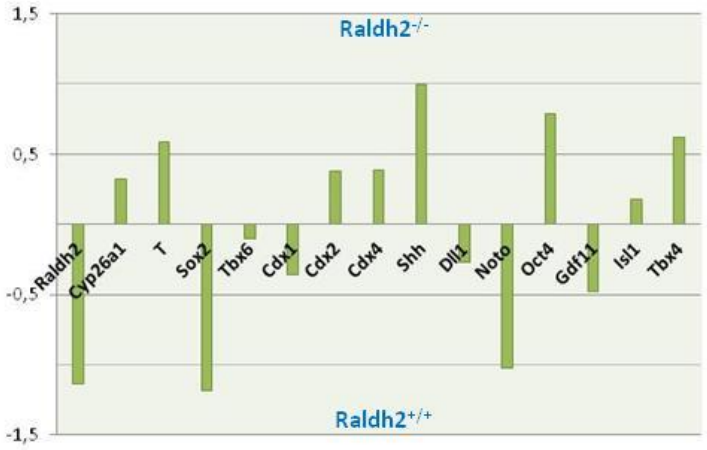
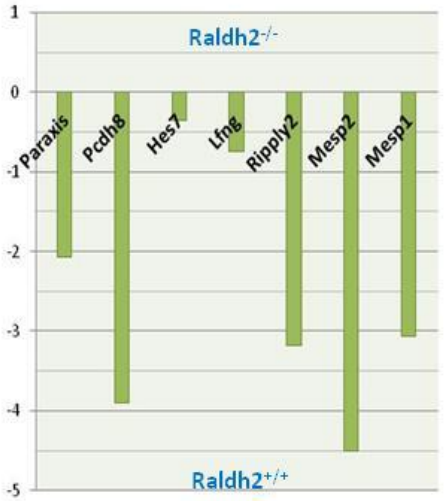
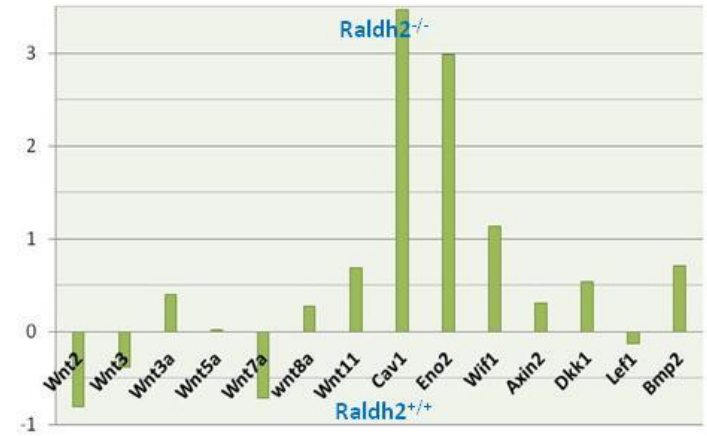
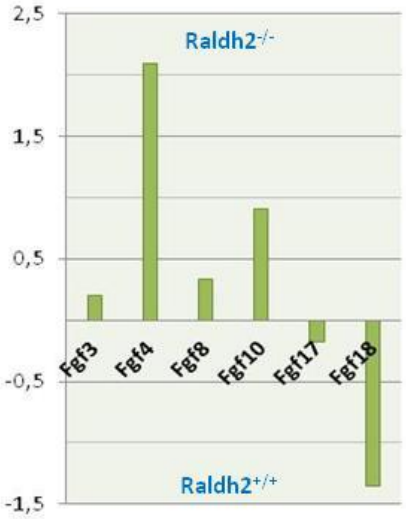


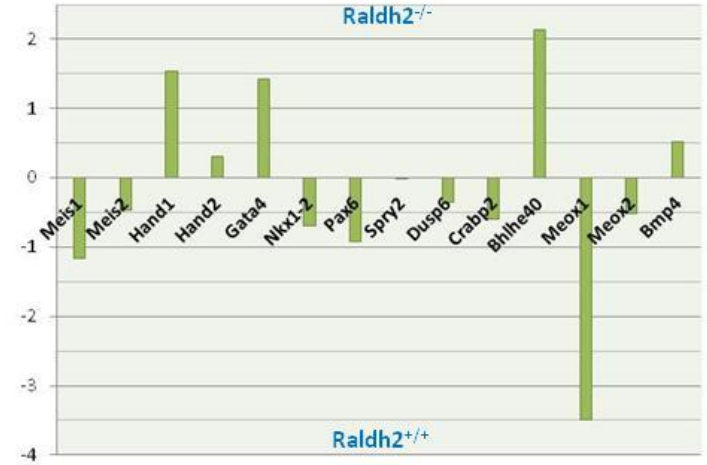
Fig. 10 – Volcano plot (t-test) using RNA-seq data, with some highlighted genes that had a high fold change in the analysis.


Raldh2-/-
Raldh2+/+
Raldh2
Cyp26a1
T
Sox2
Tbx6
Cdx1
Cdx2
Cdx4
Shh
Dll1
Noto
Oct4
Gdf11
Isl1
Tbx4


Raldh2-/-
Raldh2+/+
Paraxis
Pcdh8
Hes7
Lfng
Ripply2
Mesp2
Mesp1


Raldh2-/-
Raldh2+/+
Wnt2
Wnt3
Wnt3a
Wnt5a
Wnt7a
wnt8a
Wnt11
Cav1
Eno2
Wif1
Axin2
Dkk1
Lef1
Bmp2


Raldh2-/-
Raldh2+/+
Fgf3
Fgf4
Fgf8
Fgf10
Fgf17
Fgf18


Raldh2-/-
Raldh2+/+
Meis1
Meis2
Hand1
Hand2
Gata4
Nkx1-2
Pax6
Spry2
Dusp6
Crabp2
Bhlhe40
Meox1
Meox2
Bmp4

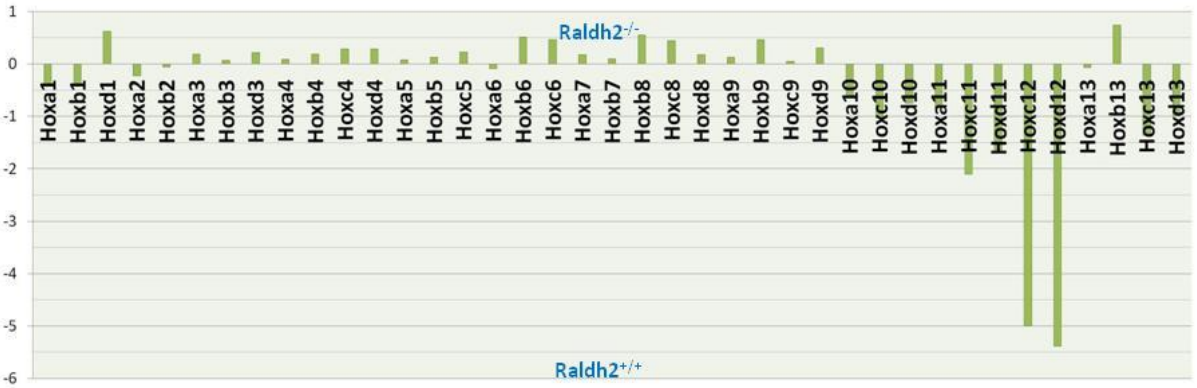

Raldh2-/-
Raldh2+/+
Hoxa1
Hoxb1
Hoxd1
Hoxa2
Hoxb2
Hoxa3
Hoxb3
Hoxd3
Hoxa4
Hoxb4
Hoxc4
Hoxd4
Hoxa5
Hoxb5
Hoxc5
Hoxa6
Hoxb6
Hoxc6
Hoxa7
Hoxb7
Hoxb8
Hoxc8
Hoxd8
Hoxa9
Hoxb9
Hoxc9
Hoxd9
Hoxa10
Hoxc10
Hoxd10
Hoxa11
Hoxc11
Hoxd11
Hoxc12
Hoxd12
Hoxa13
Hoxb13
Hoxc13
Hoxd13

Fig. 11 – Transcriptome analysis data reflecting fold change (log_2 (fold change)) **differences between selected mRNA gene expression in wild type and Raldh2$^{-/-}$ isolated tails.** Regarding Hox expression: no major changes were noticed in anterior Hox genes, the differences in posterior Hox genes can be due to developmental stage of wild type embryos and the high fold change observed in the Hox12 cluster derives from very low absolute values.

Comparison of the data obtained with the biological replicates indicated high robustness of the RNA-seq assay. Similarly, in this study, it was possible to observe in this study variations in gene expression fitting previously described gene expression experiments using other strains of *Raldh2* mutants. In particular, *Fgf8* and *Wnt3a* expression values were consistent with in situ hybridization (ISH) studies from Duester et al (2006 and 2009) [50,88]; *Fgf17*, *Fgf18*, *Wnt8a*, *Axin2*, *Cdx1*, *Cdx2* and *Cdx4* showed modifications similar to those described in Duester et al (2006)[88]; *Meis1* and *Meis2* were down-regulated, similarly to what was observed by Cunningham (2014)[89]; *T*, *Sox2*, *Sprouty2* and *Mkp3* showed expression differences similar to those in Ribes et al (2009)[90]; *Hand1* and *Fgf3* values were consistent with ISH experiments reported in Dollé et al (1999)[19]; and *Pax6*, *Bhlhe40* as well as *Crabp2* expression profiles were congruent with those obtained by Niederreither (2013)[91]. We also performed a few control tests, through ISH **(Fig.11)**, to further assess the quality of the transcriptomic data and our *Raldh2* mutation.

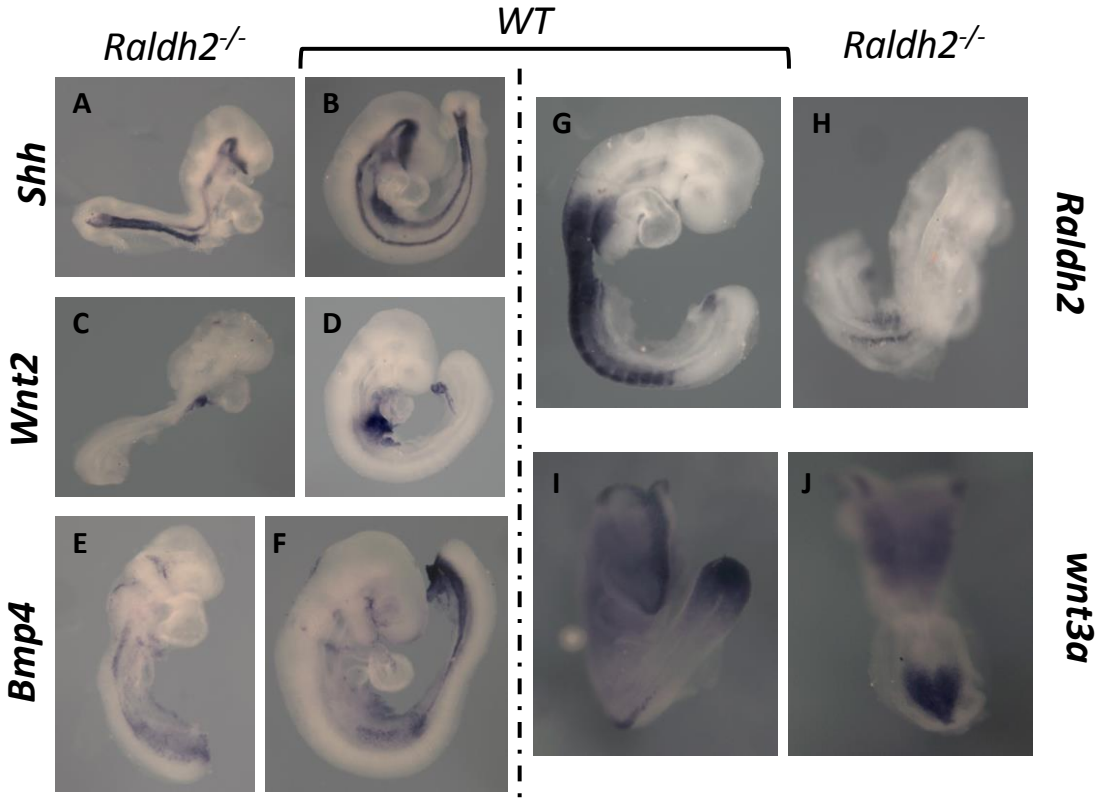


Fig. 12 – ISH analysis was congruent with data from RNA-seq. *Shh*, *Bmp4* and *Wnt3a* were up-regulated in the mutant tails (A, E, I respectively) when compared with their expression in wild type littermates (B, F, J respectively. *Wnt2* and *Raldh2* expression was downregulated in the mutants (C, G) compared to wild type tails (D, H).

Considering the previous experiments and preliminary RT-qPCR analysis **(Fig.13)** on selected genes that had a high-fold change in the RNA-seq, which seem to confirm the analysed data, we concluded that the mRNA profiles obtained provide a faithful representation of gene expression in the analyzed tissues.

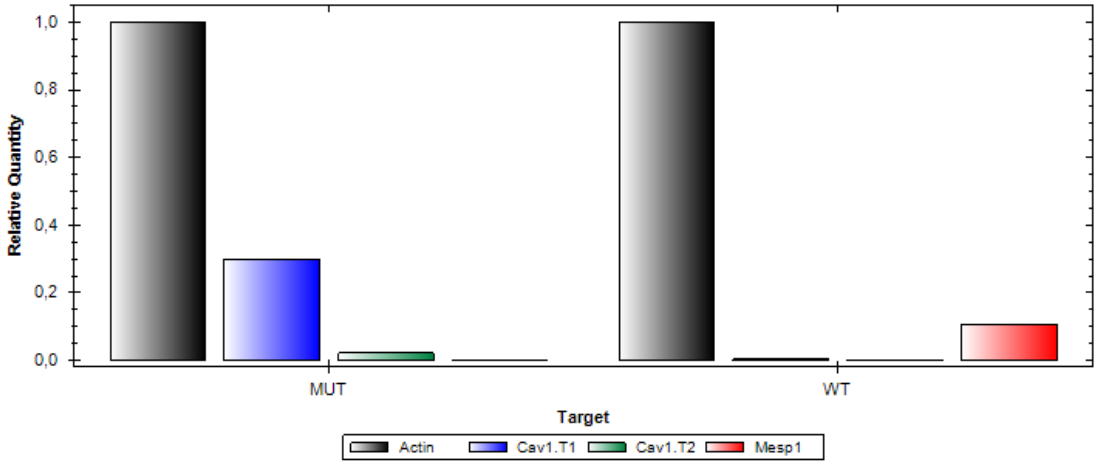


Fig.13 – RT-qPCR analysis on selected genes from the RNA-seq: *Cav1* (*Cav1.T1* and *Cav1.T2*) was found overexpressed, while *Mesp1* was downregulated in *Raldh2*$^{-/-}$.

Mesp1 was asymmetrically expressed in the *Raldh2*$^{-/-}$ presomitic mesoderm

Mesp1 is one of the genes whose expression was severely decreased in the *Raldh2* mutant tails according to our transcriptome analysis. Considering that the phenotype derived from the mutation of this gene is very similar to that of *Raldh2*$^{-/-}$ embryos (e.g. growth retardation, failure to overcome axial turning and accumulation of cells in the PS)[92,93], and because lineage tracing experiments showed that *Mesp1*-expressing cells seem to contribute to all head structures up to (and including) the forelimb buds[94], we performed ISH experiments to observe *Mesp1* gene expression in *Raldh2*$^{-/-}$ embryos. Shortly after gastrulation and during formation of the first somites, *Mesp1* expression was unchanged in the absence of embryonic RA **(Fig.14)**.

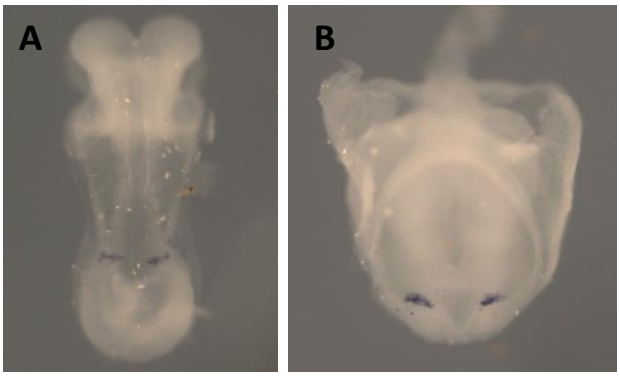


Fig. 14 – *Mesp1* expression during the first somites formation was unchanged in the absence of RA.
A – *Raldh2*$^{-/-}$ E8.0/E8.25 embryo; B – wild type E8.0/E8.25 embryo.

At later stages, when the embryo is undergoing HTT, *Mesp1* expression in *Raldh2* mutant embryos followed an asymmetric pattern similar to that observed by Vermot et al (2005)[80] for *Mesp2*. We could also observe equivalent patterns for other genes involved in segmentation[69] (**Fig.15)**.

Fig. 15 – Gene expression analysis for *Mesp1* (A, B)**, *Paraxis*** (C, D) **and *Ripply2*** (E, F) **in E8.75/E9.0 embryos**. An asymmetric expression of these genes can be observed in the *Raldh2*$^{-/-}$ embryos (A, C and E) when compared to their wild type littermates (B, D and F, respectively).

Considering the expression patterns of these genes in the PSM, it is possible that the apparent downregulation observed in the RNA-seq data resulted from the left-right somite asymmetry present in *Raldh2* mutant embryos[80], which could have led to recovery of different amounts of expressing tissue in wild type and *Raldh2* mutant embryos. Altogether, we can conclude that *Mesp1* is not responsible for the failure of *Raldh2* mutant embryos to undergo HTT.

Possible *Fgf4* overexpression in *Raldh2* mutants is not the cause for the observed embryo truncation

As referred in *Chapter I*, Fgf signalling is involved in a variety of important process during embryonic development (e.g. axial extension, somitogenesis and limb morphogenesis). Together with *Fgf8*, *Fgf4* is the main Fgf signal responsible for controlling those processes. According to Duester et al (2009), the absence of RA does not seem to affect *Fgf4* expression in the tails of E8.25 embryos. However, according to our RNA-seq data (which was performed

in a slightly different developmental stage) *Fgf4* expression in the *Raldh2* mutant embryos was higher than in wild type embryos. Therefore to understand if a deviation from the normal timing of posterior *Fgf4* down-regulation could be the cause for the mutants truncated phenotype, we tested *Fgf4* expression by ISH in *Raldh2* embryos **(Fig.16)**. The patterns that we obtained for *Fgf4* expression, and most particularly in the *Raldh2*$^{-/-}$ embryos, were not completely consistent. At E8.5, several *Raldh2*$^{-/-}$ embryos had expression in the axial stem zone, whereas from E8.75 onwards *Fgf4* expression in the posterior part of the embryo was observed only in a subset of the embryos. When this expression was present, it was restricted to a small number of cells. After 8,75 we could not detect Fgf4 expression in the tailbud of any *Raldh2*$^{-/-}$. So we concluded that even if posterior *Fgf4* expression in some Raldh2 mutant embryos was maintained longer than in wild type embryos, the low number of Fgf4 expressing cells and the timing of their expression in *Raldh2*$^{-/-}$ embryos cannot explain the truncation observed in these mutants and therefore Fgf4 does not seem to play a major role in the HTT.

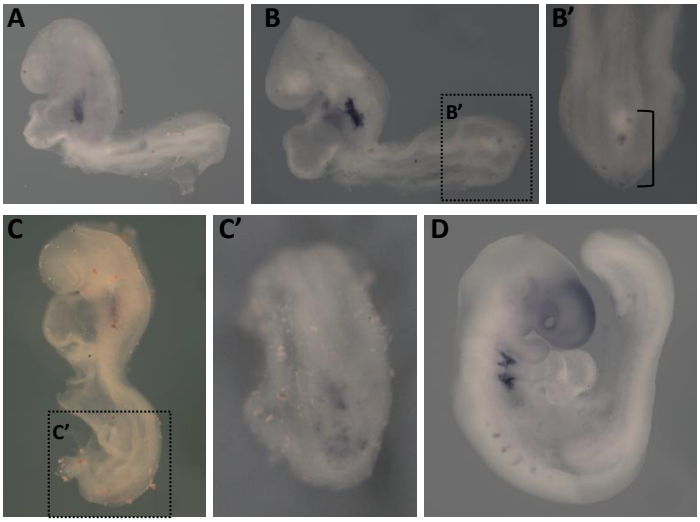


Fig. 16 – Fgf4 expression was observed in the PSM of only some Raldh2 mutants: A – Example of E8.75 Raldh2$^{-/-}$ were no expression was detected; B and C – mutants embryos (E8.75 and E9.0, respectively) were Fgf4 was detected in the PSM. D – Wild type E9.0 were Fgf4expression was absent in the PSM.

Canonical Wnt signalling activity was reduced in the tails of Raldh2 mutant embryos

In the RNA-seq data sets we could observe differential expression of several genes involved in the Wnt signalling in the *Raldh2* mutant tails (e.g. *Wnt3a, Cav1 and Wif1*). Considering the role of Wnt signalling during axial extension, particularly the interactions and molecular functions of *Wnt3* and *Wnt3a* (fully described in Chapter I) we decided to explore a possible RA-mediated change in Wnt signalling at the time of the HTT. For that we used the TopGal Wnt reporter transgenic mice in order to observe the activity of the canonical Wnt pathway **(Fig.17)**. In the tail of *Raldh2*$^{-/-}$ embryos, we found that this activity was reduced when compared to wild type embryos, which suggests that a specific canonical Wnt activity indirectly depends on the presence of retinoic acid. This result was interesting since *Wnt3a* expression

was higher in *Raldh2* mutant tissues than in wild type controls, indicating that the lower Wnt activity in the *Raldh2*$^{-/-}$ did not derive from lower signal production. Therefore, we decided to test the involvement of some Wnt inhibitors that were differentially expressed in the *Raldh2*$^{-/-}$ mutants according to the RNA-seq data (e.g. *Dkk1*, *Cav1* and *Eno2*) in the molecular mechanism that controls the HTT.

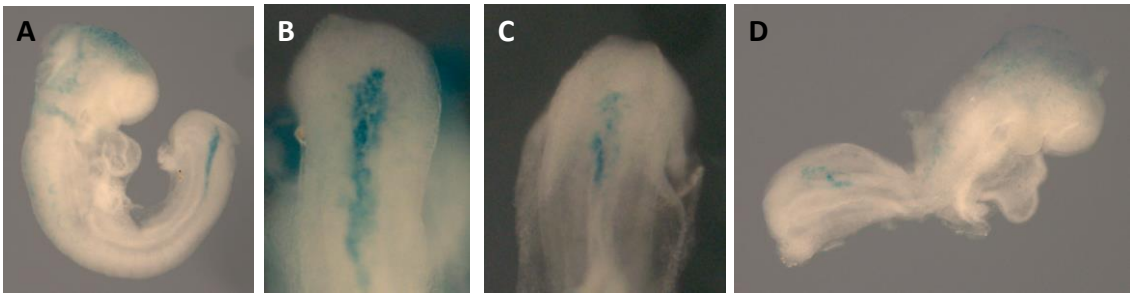


Fig. 17 – Wnt canonical activity using TopGal reporter in wild type (A) **and in *Raldh2*$^{-/-}$** (D)**.** This activity is reduced in the tails of *Raldh2*$^{-/-}$ (C) when compared to what occurs in wild type tails (B).

Dkk1 expression does not seem to be altered in the absence of retinoic acid

Dickkopf-1 is one of the Wnt inhibitors presumably overexpressed in *Raldh2*$^{-/-}$ tails according to RNA-seq data. The genetic knockout of *Dkk1* showed that it is essential for the induction of anterior head structures and that an antagonist interaction between Dkk1 and Wnt3 is crucial not only for head but also for trunk morphogenesis[95,96]. Since in normal embryos its expression seems to decay from E8.25/E8.5 to E9.5[97–99] we decided to investigate if the apparent higher *Dkk1* expression observed in *Raldh2*$^{-/-}$ embryos could be involved in the lower Wnt signalling observed in these embryos. *Dkk1* expression in *Raldh2* mutants reproduced the patterns described in Lewis et al (2007) for E8.25/E8.5 wild type embryos. Also, and quite surprisingly, we detected *Dkk1* expression in the tails of both E9.5 and E10.5 wild type embryos, indicating that *Dkk1* expression in the tailbud is compatible with normal axial extension. Therefore, our ISH experiments **(Fig.18)** do not support the requirement of a RA-dependent down-regulation of *Dkk1* for HTT. A possible source for the differences in *Dkk1* expression observed in the RNA-seq datasets between wild type and *Raldh2*$^{-/-}$ tails might be differences in cycling behaviour of *Dkk1* expression[100], although evaluation of this hypothesis requires a direct experimental approach.

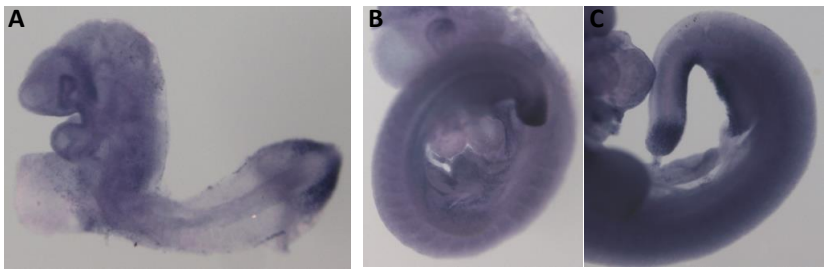


Fig. 18 - *Dkk1* expression in the PSM of *Raldh2*$^{-/-}$ (A) **and wild type littermates** (B) **around E9.0 and in wild type E9.5 embryos** (C)**.**

Is Caveolin 1 involved in the head to trunk transition?

Cav1 has been suggested to block the canonical Wnt signalling pathway by retaining β-catenin into the cell's membrane[101]. In our RNA-seq dataset, *Cav1* came out as a gene strongly up-regulated in the tails of *Raldh2* mutant embryos. Although *Cav1* inactivation does not lead to embryonic lethality[102], it is still possible that its overexpression (as seen in the *Raldh2* mutants) could result in serious damages to the embryo due to inhibition of Wnt signalling. In wild type embryos *Cav1* transcripts were observed in a small domain in the anterior PSM at E8.25/E8.5. *Cav1* expression seemed more widespread in E8.75 *Raldh2*$^{-/-}$ embryos, where it presented a scattered pattern (in some embryos more expression was noticed than in others) **(Fig.19)**. This difference in *Cav1* expression between wild type and *Raldh2*$^{-/-}$ embryos led us to perform a transgenic experiment to evaluate if overexpression of *Cav1* in the axial progenitors could block the HTT.

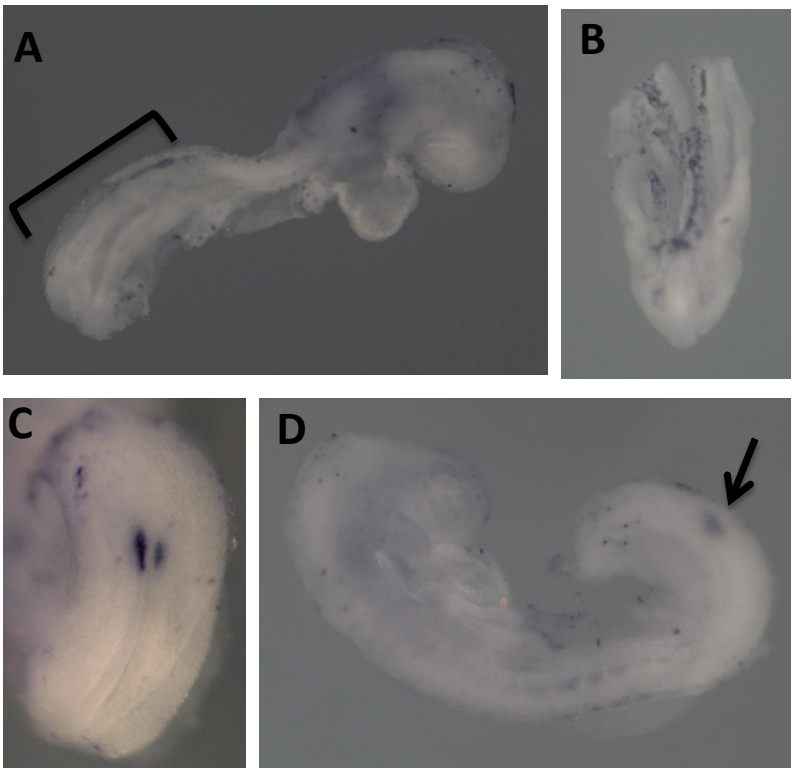


Fig. 19 – *Cav1* expression was detected in the PSM of some E8.75 *Raldh2*$^{-/-}$ (A and B) **and in E8.25** (C) **and E8.5** (D) **wild type embryos.**

Cav1 has two transcript variants, *Cav1T1* and *Cav1T2* representing the longer and shorter isoforms, respectively. We produced transgenic embryos expressing each of the *Cav1* transcript variants under the regulation of a *Cdx2* promoter. At E9.5 *Cdx2.Cav1T1* embryos had no visible phenotype **(Fig.20)**, whereas we were unable to obtain *Cdx2.Cav1T2* embryos at E9.5 and E8.25 **(Table 7)**. Interestingly, it seems that only the smaller transcript variant of Cav1 codes for the complete caveolin1 protein. Therefore, it is possible that our inability to recover

Cdx2.Cav1T2 embryos could derive from a negative effect of the protein encoded by *Cav1T2* at earlier developmental times, which might have resulted from inhibition canonical Wnt signalling before gastrulation. Indeed, the *Cdx2* enhancer used in these experiments contained an element driving expression at early developmental stages. Analysis of *Cdx.Cav1T2* at earlier developmental times, together with additional transgenic experiments using other enhancers will be required to determine if this is indeed the case.

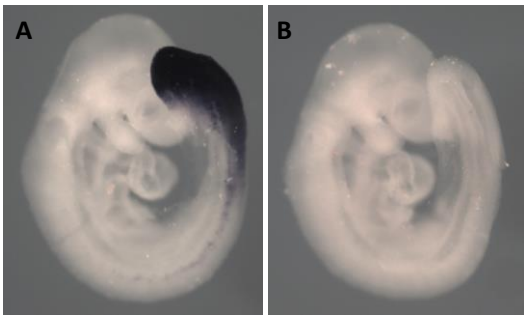


Fig. 20 – *Cav1* RNA in situ gene expression assay in *Cdx2.Cav1T1* transgenic embryos (A) **and in their wild type littermates** (B)

Microinjection	Stage	Embryos / Genotype	Reabsorptions
1st using 5 females	~E9.5	13 – all WT	26
2nd using 4 females	~E8.25	16 – all WT	11

Table 7: Cdx2.Cav1T2 transgenic embryos genotype

Retinoic acid and the axial progenitor cells

A hallmark of the transition from head to trunk development is the appearance of the bipotent NMPs that drive elongation of the neural tube and somites. These NMPs are loosely defined as Sox2+/T+[58]. Considering that *Raldh2* mutant embryos are blocked at the HTT stage we decided to explore formation of the NMPs in these embryos. The RNA-seq data suggested *T* up-regulation and *Sox2* down-regulation in the *Raldh2*$^{-/-}$ tails, which was congruent with the findings reported by Ribes et al (2009) for E8.5 embryos. ISH experiments on E8.75/E9.0 embryos confirmed these data **(Fig.21)**. Also, a close analysis of these data revealed an accumulation of *T* expressing cells in tip of the tail, whereas *Sox2* appeared to be down-regulated near the PS, in the region where the NMPs are localized at this stage. These observations suggested a loss of *T* and *Sox2* co-localization, which at this stage of development indicates the presence and position of the axial progenitor cells. To further explore this finding we decided to perform immunohistochemistry experiments using antibodies for both proteins in order to characterize their localization in the tails of *Raldh2*$^{-/-}$ embryos. In preliminary experiments (data not shown) we could observe that T and Sox2 co-localized at the protein level in the *Raldh2*$^{-/-}$ tails, consistent with the presence of axial progenitor cells. However, although the data is limited, it is possible that the number of double positive cells was reduced compared to

wild type embryos. Further experiments will be required to perform a proper quantitative evaluation of the number of T+/Sox2+ progenitors in the *Raldh2* mutant embryos.

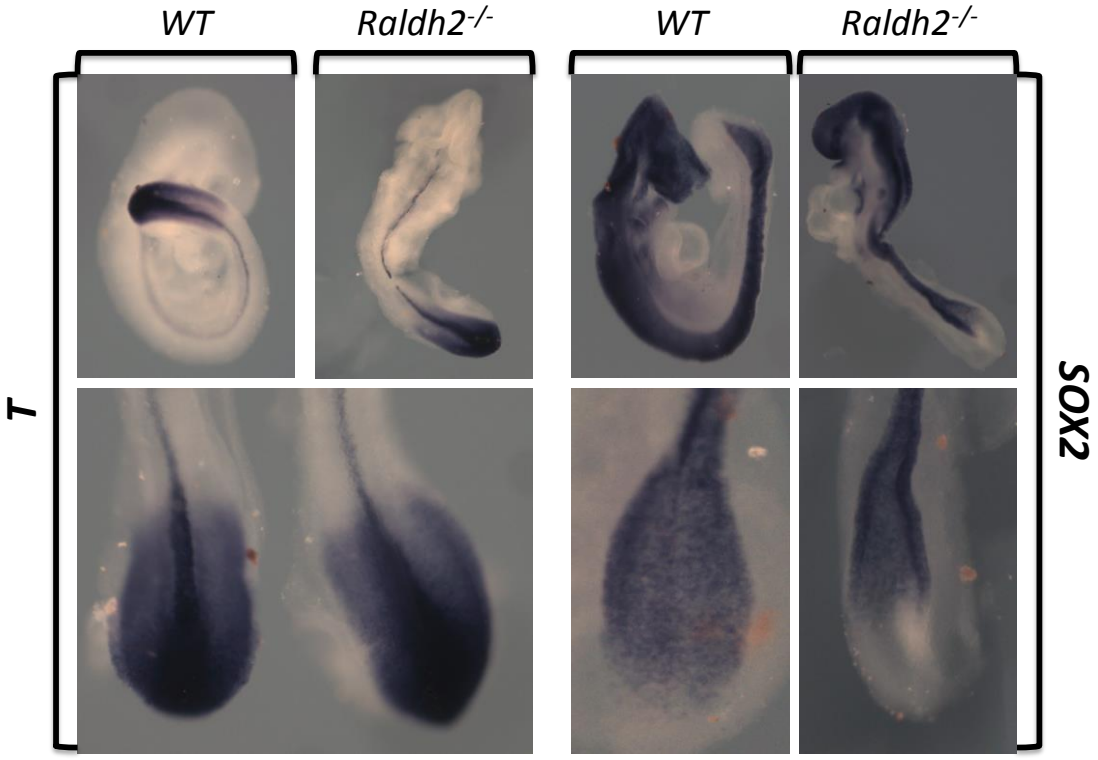


Fig 21. Accumulation of *T* expressing cells in the tails of *Raldh2*$^{-/-}$ E8.75 embryos, where *Sox2* was downregulated near the PS.

Tbx5 expression can be observed in *Raldh2* mutants

Finally, in the course of the previous experiments, we noticed the presence of a protuberance symmetrically located near the heart on both sides of *Raldh2*$^{-/-}$ embryos at ~E9.5, resembling small forelimb buds. ISH experiments with *Tbx5*, the first known marker of the forelimb buds[86], confirmed the identity of those protuberances **(Fig.22)**. This was surprising because *Raldh2* mutants have been reported to be unable to induce forelimbs[76]. To better understand how RA controls the HTT, further studies will be necessary to observe the timing of forelimb bud induction and the position it assumes relative to the forming somites in *Raldh2*$^{-/-}$ embryos.

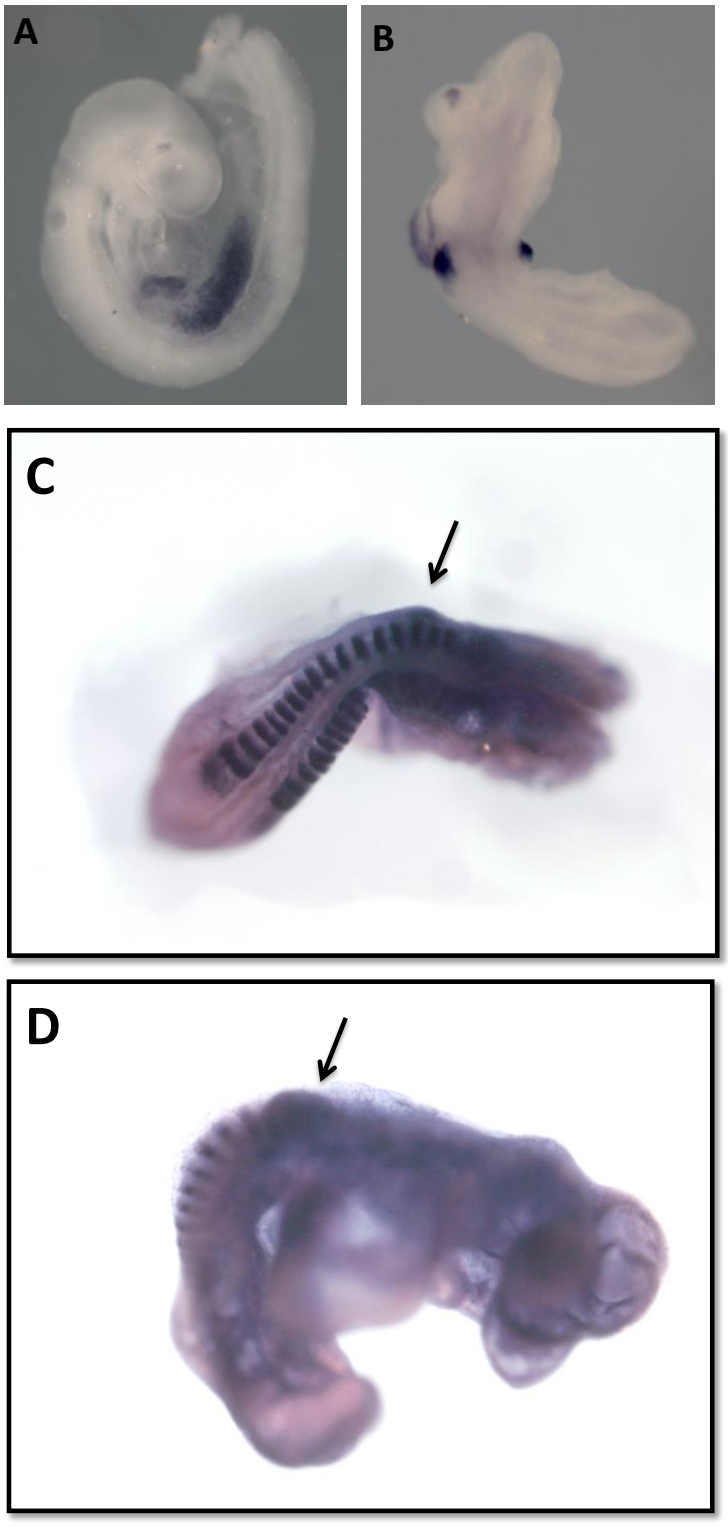


Fig. 22 – *Tbx5* expression domain in the presumptive forelimb bud domain of *Raldh2* mutant embryos (B-D) and wild type littermates (A). Analysis was performed by ISH using a probe for *Tbx5* alone (A and B) or combined with a probe for *Uncx4.1* (C and D).

Discussion

Retinoic acid signalling seems to set the time for the head to trunk transition, possibly by controlling axial progenitor activity

Raldh2 mutant embryos show a developmental block around E8.25, the time at which the HTT occurs[19]. The finding that this phenotype can be rescued by exogenous administration of RA coincident with the time of the developmental arrest[20], suggests that RA signalling is crucial for the embryo to undergo HTT. In the embryo, the levels of RA result from the balance between its synthesis by Raldh2 and its degradation by Cyp26a1[24]. *Cyp26a1* mutant embryos are truncated at the lumbar level, indicating that removal of RA is required for the TTT[10]. Interestingly, however, *Cyp26a1* mutant embryos also have a posterior transformation of cervical vertebra (C) 5 into C6 and of C7 into a thoracic vertebra acquiring an extra pair of ribs[23], which could result from an earlier HTT. Because in the absence of *Cyp26a1*, which starts to be expressed around E8.25[23,24], the embryo is expected to accumulate an excess of RA in the progenitor-containing area (after gastrulation this gene is expressed by the node), it is possible that the premature HTT observed in these embryos results from higher levels of RA activity. This effect would then be complementary to the inability of *Raldh2* mutants to undergo proper HTT.

One of the key features of the HTT is the appearance of NMPs that drive elongation of the neural tube and paraxial mesoderm in the trunk and tail areas. These progenitors are thought to be positive for T and Sox2[58]. We have found that in the caudal end of *Raldh2*$^{-/-}$ mutant embryos the area containing both transcripts, as determined by ISH, was reduced when compared to wild type embryos. It is then possible that in *Raldh2*$^{-/-}$ the production of axial progenitors is reduced and/or its activity changed due to the inability to undergo HTT. Using an immunohistochemistry approach we could identify tail cells expressing both T and Sox2 proteins, indicating the possible existence of NMPs in *Raldh2*$^{-/-}$ embryos. These experiments are still too preliminary to determine if the number of double T/Sox2 positive cells is reduced in *Raldh2*$^{-/-}$ when compared to their wild type littermates. Also, experiments from Mallo's laboratory indicate that the Sox2 protein is stable for much longer than the corresponding transcript. It is therefore possible that as a consequence of the reduced Sox2 transcripts, the levels of Sox2 protein will be progressively reduced in the progenitor area as development proceeds, eventually resulting in exhaustion of NMPs. In the future we will continue to address experimentally these hypotheses to understand how RA affects the activity of the NMPs during this particular developmental stage, thus controlling the time at which the HTT occurs.

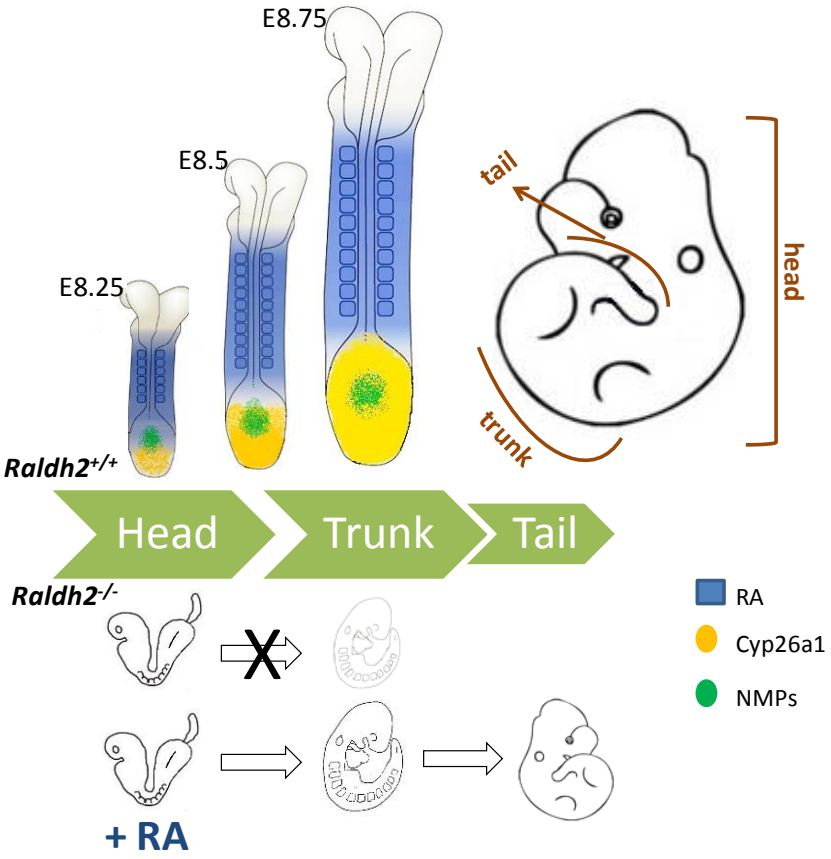


Fig. 23 – Requirement of RA during head and trunk formation. RA activity is necessary, between E8.25/E8.5, in axial progenitor cells in order to switch from producing head to trunk tissues in the embryo. Cyp26a1, which starts to be expressed from this stage onwards protects the NMPs from RA activity.

Does retinoic acid control the head to trunk transition by promoting a change in the Wnt signalling?

The mechanism by which RA promotes the HTT is still unclear. Considering the *Wnt3* and *Wnt3a* expression patterns and their differential roles in head and trunk formation (described in *Chapter I*), it is possible that RA activity is necessary to change the specific requirements for Wnt signalling during the HTT. Such change in the Wnt signalling pathway at this stage is supported by the finding that stabilization of Axin2, while leading to a decrease of Wnt activity during gastrulation as expected from its known negative effects on Wnt signalling, it resulted in higher Wnt activity in the PS of E8.5 embryos[103]. Interestingly, when investigated what canonical Wnt activity in E8.5 *Raldh2* embryos using a transgenic reporter assay, we found a decrease in Wnt activity in the tailbuds of *Raldh2* mutant embryos, thus consistent with a hypothetical role of RA in the apparent change in Wnt signalling during HTT. These findings are, however, contradictory with the observations indicating that *Wnt3a* expression is apparently up-regulated in the *Raldh2*$^{-/-}$ tails relative to their wild type littermates. It is thus possible that the change in Wnt activity is not exerted at the level of the signal itself but results from a regulatory event affecting its ability to activate signalling in the recipient cells. Analysis

of our RNA-seq data indicated that several inhibitors of the β-catenin pathway, including *Cav1*, *Dkk1*, *Eno2*, *Axin2*, *Wif1* and others, were present at higher levels in RA-negative embryos. So far, our functional data was not sufficient to evaluate the contribution of these factors to the hypothetical switch in the Wnt pathway during HHT, but we are performing additional experiments to test it. For instance, we intend to chemically stabilize Axin2 in the *Raldh2*$^{-/-}$ and then observe what happens to canonical Wnt activity. We are also in the process of performing gain and loss of function experiments in order to observe if these genes could alter Wnt signalling during the HTT and to observe the effect that the rescue of the *Raldh2*$^{-/-}$ phenotype with RA treatments has on the expression of these genes. Also, it will be interesting to understand the regulation of these genes by RA signalling.

A retinoic acid-dependent dual origin for forelimb bud induction

A surprising observation we made in this work was the identification of *Tbx5*-positive protruding structures next to the developing heart in *Raldh2*$^{-/-}$ embryos. Both the position of these structures and their expression of *Tbx5* identify them as forelimb buds. This was striking considering that *Raldh2*$^{-/-}$ embryos have been reported to be unable to induce forelimbs[19,20]. Interestingly, this *Tbx5*-positve forelimb domain was very similar to what was observed in *Rdh10*$^{trex/trex}$ embryos, which undergo HTT forming both trunk and tail structures (including hindlimbs)[18]. Also, a small forelimb bud develops in *Raldh2*$^{-/-}$ that had been rescued by RA, in a dose-dependent manner[78]. As these embryos, similarly to *Rdh10*$^{trex/trex}$, undergo HTT, it is possible that the absence of forelimbs in *Raldh2*$^{-/-}$ embryos has a dual origin, one derived from the inability to undergo HTT and another from a direct requirement of RA signalling to continue the induction and initiation process necessary for complete forelimb bud development. Similarly to the connection between hindlimb induction and TTT, our present results now suggest a link between HTT and forelimb induction.

Acknowledgments/Contributions

André Dias performed most of the experiments reported in this chapter with some contributions from members of the Mallo lab and collaborators. In particular, Moisés Mallo and Ana Nóvoa generated the *Raldh2* mutants, Ana Nóvoa performed all microinjections to generate the transgenic embryos, the sequencing of the RNA (RNA-seq) was part of a collaboration with the core facility at the EMBL and Rita Aires helped with the immunohistochemistry experiment.

Chapter III

A novel approach to the *Gdf11$^{-/-}$::T-streakCre^{ERT}#47$^{+/0}$* problem:

The combination effect...

The *Gdf11*$^{-/-}$*::ROSA26R-ß-gal*$^{+/0}$*::T-streakCreER*^{T}*#47*$^{+/0}$ apparent embryonic lethality

Gdf11 signalling plays an essential role in axial patterning during vertebrate development. It is involved in a variety of processes, including the control of the trunk to tail transition, where it regulates the fate of the NMPs when they migrate from the epiblast into the chordo-neural hinge (CNH). The involvement of *Gdf11* signalling in this process can be illustrated by the observation that some *Gdf11*$^{-/-}$ embryos exhibit split tails, where *T* expression is also segregated, due to the incorrect reallocation of NMPs into the CNH[9,10,25].

In the course of experiments conducted by Arnon Jurberg with the objective of understanding how Gdf11 signalling modulates the activity of these progenitors, an experimental approach was developed consisting in the introduction into the *Gdf11* mutant background a transgene expressing *Cre*^{ERT} under the control of a PS-specific regulatory element of *T* (*T-streakCre*^{ERT}) to activate the *ROSA26R-ß-gal* reporter in the nascent axial progenitors. Results obtained from these experiments indicated that *Gdf11*$^{-/-}$*::ROSA26R-ß-gal*$^{+/0}$*::T-streakCreER*^{T}*#47*$^{+/0}$ embryos apparently died before E10.5, independently of whether or not they had been treated with tamoxifen **(Table 8)**. Further preliminary studies also indicated a very early lethality when the *T-streakCreER*^{T}*#47* transgene was present in *Gdf11*$^{-/-}$ embryos[8].

Gdf11$^{+/-}$	viable and fertile	Gdf11$^{-/-}$	die at birth
R26R-β-gal$^{+/+}$	viable and fertile	Gdf11$^{-/-}$::T-streak-CreER^{T}#47$^{+/0}$	possibly die at E7.5
T-streak-CreER^{T}#47$^{+/+}$	viable and fertile	T-streak-CreER^{T}#49$^{+/+}$	viable and fertile
Gdf11$^{+/-}$::Rosa26R-β-gal$^{+/+}$	viable and fertile	Gdf11$^{+/-}$::T-streak-CreER^{T}#49$^{+/+}$	viable and fertile
Gdf11$^{+/-}$::T-streak-CreER^{T}#47$^{+/+}$	viable and fertile	Gdf11$^{-/-}$::T-streak-CreER^{T}#49$^{+/0}$	?

. Gdf11 genotyping in the *R26R-β-gal*$^{+/0}$*;T-streak-CreER*$^{T+/0}$ #47 background[1]. ~E10.5

Genotype	*Number of embryos obtained*	*Frequency (%)*	
		Observed	Expected
Gdf11$^{+/+}$	68	44.2	25
Gdf11$^{+/-}$	86	55.8	50
Gdf11$^{-/-}$	0	0	25
Total	154	100	100

[1]The fit to Mendelian expectation was tested with a chi-square test: $\chi^2 = 62.16$, degrees of freedom = 2, $P<0.0001$.

Table 8 – Genetic background and crossings performed to obtain *Gdf11*$^{-/-}$*::ROSA26R-ß-gal*$^{+/0}$*::T-streakCre*^{ERT}*#47*$^{+/0}$ embryos. No such embryos were found at E10.5 according to Mendelian expectations (adapted from Jurberg PhD Thesis)

To understand the origin of the lethality of *Gdf11*$^{-/-}$*::T-streakCre*^{ERT} mice, Jurberg determined the integration site of the *T-streakCre*^{ERT} transgene in the transgenic line *#47* (the one used in his experiments). He found that it was integrated on chromosome 14, between *Pcdh8* and *Olfm4*. Two additional uncharacterized putative protein-coding transcripts (Gm6999 and Gm10845) were also present in this region. To see if the transgene integration disrupted regulation of these genes, Rita Aires performed ISH with a *Pcdh8* probe. In these experiments, she observed a possible misexpression of this gene in the posterior region of *T-streakCre*^{ERT} *#47*$^{+/+}$ embryos at E7.5 suggesting that this misregulation could be involved in the early lethality of *Gdf11*$^{-/-}$*::T-streakCre*^{ERT} embryos[8,87].

Is the *Gdf11*$^{-/-}$*::T-streakCre*^{ERT}*#47*$^{+/0}$ lethality due to a combination effect?

The specific aim of this project was to further explore the origin of the apparent early lethality of *Gdf11*$^{-/-}$ *::T-streakCre*^{ERT}*#47*$^{+/0}$ embryos, under the hypothesis of a possible combination effect resulting from the *Gdf11* inactivation plus the disruption of a particular conserved region (involving *Olfm4* and *Pcdh8*), caused by the insertion of the transgene in the *#47* line.

Material and methods

[§]Solutions detailed information is present in **Supplementary Information II**

*All manufacturer protocols are present in **Supplementary Information III**

[#]Plasmid maps are shown in **Supplementary Information IV**

Mouse strains

The *Gdf11*$^{+/-}$ strain was described in McPherron (1999) and was kindly provided by the authors. It was maintained at the IGC Rodent Facility on a C57BL/6J background. Transgenic mouse lines *#47* and *#49* were created by pronuclear injection of the *T-streak-Cre*^{ERT} construct (provided by A. Gossler and B. Herrmann) in the FVB/N genetic background. The *ROSA26R-β-gal* reporter was described in Soriano (1999) and was purchased from the Jackson laboratories.

Housing conditions, mating and embryo harvesting were similar to what was described in *Chapter II (material and methods section)*

Oligos

Primers for	Forward	Reverse
Gdf11 $_{Mut}$	GGATCGGCCATTGAACAAGATG	GAGCAAGGTGAGATGACAGGAG
Gdf1 $_{WT1}$	GAGTCCCGCTGCTGCCGATATCC	TAGAGCATGTTGATTGGGGACAT
Gdf11 $_{WT2}$	CTGCTGCACCCCTACCAAGATG	CCACTGTAGCCCACAACTTAGGAG
T-streak-CreER ^{T}	CGAGTGATGAGGTTCGCAAG	CACCAGCTTGCATGATCT
Olfm4 $_{CLON}$	CGCTCGAGGCGCTCCTTCTGTGATTCAGC	CCGCGGCCGCTGAAGGCATTCGAAACAGATGG

Table 9 –Primers used for PCR.

Genotyping and PCR conditions

The process for genotyping mice (using tail biopsies) and embryos (using yolk sacs) was described in Chapter II (material and methods section). Specific PCR conditions are described on **Table 10**.

<table>
<tr><th colspan="3">Oligos</th><th colspan="2">PCR conditions</th></tr>
<tr><td colspan="3">$Gdf11_{Mut}$</td><td rowspan="4">95°C for 5 min</td><td>34cycles (95°C for 45 sec, 58°C for 1 min and 72°C for 2 min)</td><td rowspan="4">72°C for 10 min</td></tr>
<tr><td rowspan="3">$Gdf11_{WT1}$</td><td colspan="2">$Gdf11_{Mut}$</td><td rowspan="2">34cycles (95°C for 45 sec, 55°C for 1 min and 72°C for 2 min)</td></tr>
<tr><td colspan="2">+</td></tr>
<tr><td colspan="2">$Gdf11_{WT1}$</td></tr>
<tr><td>$Gdf11_{WT2}$</td><td>$T\text{-}streak\text{-}CreER^{T}$</td><td>$Olfm4_{CLON}$</td><td>34cycles (95°C for 45 sec, 62°C for 1 min and 72°C for 2 min)</td></tr>
</table>

Table 10 – Specific PCR conditions

Olfm4 probe

500 bps of the *Olfm4* coding sequence were amplified from mouse genomic DNA (from E18.5 embryo intestines) with the $Olfm4_{CLON}$ primers (**Table 9**). The PCR-amplified fragment was then purified using the *MinElute** and *Qiaquick* Kits* (both from *QIAGEN*) and cloned into the XhoI and NotI sites of pBluescript II KS[#]. Molecular cloning, MINI- and MIDI-scale plasmid preparations, as well as the sequencing reactions, were performed as described in *Chapter II (material and methods section)*. For in vitro transcription, the plasmid was linearized with XhoI and the probe was synthesized with T7 RNA polymerase.

Other probes

All the other used probes were available in the lab and had been synthesized by in vitro transcription.

In situ Hybridization

The protocol used for ISH was described in *Chapter II (material and methods section)*. For earlier stages than E8.0, after the rehydration step, embryos were placed in a box with grids (the box was previously washed with NaOH to remove RNases), where the whole ISH procedure was performed. These embryos were genotyped after ISH was completed. For this, after taking pictures of the stained embryos, they were incubated in PBND[§] with ¼ of Proteinase K according to their stage. After inactivation, about 5 µL of DNA was used in a PCR reaction with 40 cycles and with 5 mM of $MgCl_2$.

Results

Gdf11$^{-/-}$*T-streak-Cre*^{ERT}*#49*$^{+/0}$ embryos were alive at E10.5

The first step to understand the origin of the apparent early lethality of *Gdf11*$^{-/-}$ *::T-streakCre*^{ERT}*#47*$^{+/0}$ embryos was to introduce the *Gdf11* mutation into another line of the *T-streakCre*^{ERT} transgene (*#49*). When crossing *Gdf11*$^{+/-}$*::T-streak-Cre*^{ERT}*#49*$^{+/+}$ females with *Gdf11*$^{+/-}$ males we could recover *Gdf11*$^{-/-}$*::T-streak-Cre*^{ERT}*#49*$^{+/0}$ embryos at E9.5 and E10.5. These embryos were phenotypically similar to *Gdf11*$^{-/-}$ embryos, indicating that the *T-streak-CreER*^{T}*#49* transgene did not induce early lethality to *Gdf11*$^{-/-}$ embryos or modified their phenotypic characteristics. Therefore, it is not the transgene per se but the position where it was integrated in the genome what might have contributed to the apparent lethality of *Gdf11*$^{-/-}$*::T-streak-Cre*^{ERT}*#47*$^{+/0}$ embryos.

Phenotypic characterization of *Gdf11*$^{-/-}$*::T-streak-Cre*^{ERT}*#47*$^{+/0}$ embryos

A phenotypic characterization of *Gdf11*$^{-/-}$*::T-streak-Cre*^{ERT}*#47*$^{+/0}$ embryos was then performed to determine the developmental stage and cause(s) of their death. Since *Gdf11* starts to be expressed around E7.5[104], we isolated E7.5 embryos from intercrosses between *Gdf11*$^{+/-}$ males and *Gdf11*$^{+/-}$*::T-streak-Cre*^{ERT}*#47*$^{+/+}$ females. Some of these embryos looked morphologically different from their littermates, displaying a softer texture (possibly because they were in the process of being reabsorbed). All these apparently affected embryos were *Gdf11*$^{-/-}$*::T-streak-Cre*^{ERT}*#47*$^{+/0}$ **(Fig.24)**, which lead us to speculate that indeed *Gdf11*$^{-/-}$*::T-streak-Cre*^{ERT}*#47*$^{+/0}$ embryos might die around E7.5, possibly due to a problem during gastrulation.

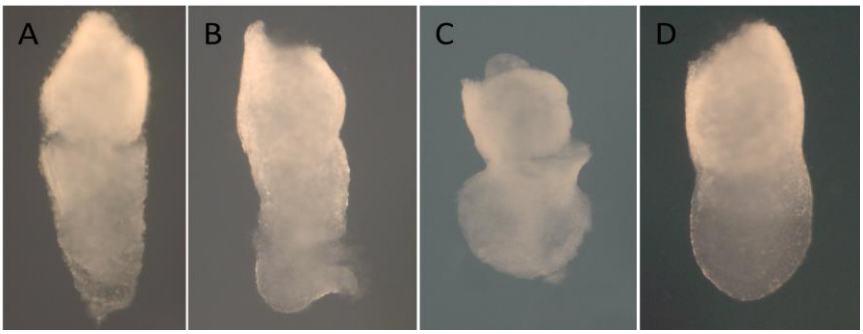
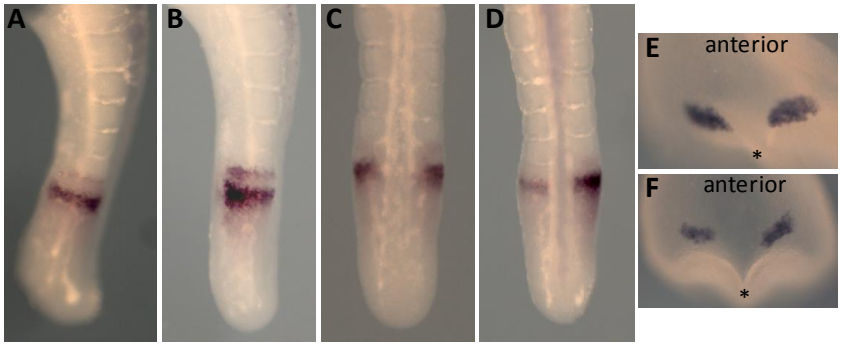


Fig. 24 – Phenotypic characterization of *Gdf11*$^{-/-}$*::T-streak-Cre*^{ERT}*#47*$^{+/0}$ (A, B and C) **and *Gdf11*$^{+/+}$*::T-streak-Cre*^{ERT}*#47*$^{+/0}$** (D) **embryos.**

Molecular characterization of *T-streak-Cre*^{ERT}*#47* embryos

To address the hypothesis of a possible combination effect where the *T-streak-CreER*^{T}*#47* transgene might strengthen the *Gdf11*$^{-/-}$ phenotype, we performed several gene expression studies to characterize molecularly the embryos from the line *#47*. Because in this line the *T-streak-Cre*^{ERT} was inserted in a genomic region next to the *Pcdh8* and *Olfm4* genes, we tested if the presence of this transgene might have affected their expression. In several

independent experiments, *Olfm4* expression was not detected at E7.5 in neither WT nor *T-streak-Cre^{ERT}#47$^{+/+}$* embryos. At this stage, a very weak *Pcdh8* expression was observed *(data not shown)* in the PS of some but not all analysed embryos (*T-streak-Cre^{ERT}#47$^{+/+}$* and WT). Although ISH does not allow proper quantification of gene expression, the observed trend suggested a possible over-expression of this gene in E7.5 *T-streak-Cre^{ERT}#47* embryos, in agreement with previous preliminary observations made by Rita Aires. Only at the head-fold stage *Pcdh8* expression was clearly detected near the node **(Fig. 25)**, fitting with previously described patterns[105]. However no abnormalities in expression were noticed when transgenic embryos were compared to wild type littermates. Through ISH we also observed patterns compatible with an oscillating behaviour of *Pcdh8* expression in the PSM of E10.5 embryos. Again, no differences were found between the *T-streak-Cre^{ERT}#47$^{+/+}$* and wild type embryos **(Fig.25)**. These results indicate that if the insertion of the *T-streak-Cre^{ERT}* transgene affected *Pcdh8* expression in the #47 mice line this happened only at very early developmental stages.



(Fig 25) – In situ hybridizations in *T-streak-Cre^{ERT}#47$^{+/+}$* and WT embryos. Lateral (A, B) and frontal (C, D) views of *Pcdh8* expression in *T-streak-Cre^{ERT}#47* (A, C) and wild type tails (B, D) of E10.5 embryos. In E8.0 embryos, the expression domain can be observed near the node (*), both in *T-streak-Cre^{ERT}#47* (E) and wild type (F) embryos.

Gdf11$^{-/-}$::T-streak-Cre^{ERT}#47$^{+/0}$ embryos gastrulate correctly

Considering the possible alteration of *Pcdh8* expression in *T-streak-Cre^{ERT}#47* gastrulating embryos and the preliminary results, we searched for patterning defects at this stage in the *Gdf11$^{-/-}$::T-streak-Cre^{ERT}#47$^{+/0}$* embryos. In gastrulating *Xenopus* embryos, *Pcdh8* was found to inhibit the canonical Wnt signalling[106], which in mouse embryos plays a role in mesoderm formation in part by regulating of *Cripto* (a *Nodal* co-receptor) expression[32]. During gastrulation *Nodal* signalling pathway is controlled by RA[7]. The observation that RA levels are altered in the tails of E9,5 *Gdf11$^{-/-}$* embryos (because during the trunk to tail transition *Cyp26a1* activity is reduced in these tissues[10,25]), led us to hypothesize that if RA signalling is reduced in E7.5 *Gdf11$^{-/-}$::T-streak-Cre^{ERT}#47$^{+/0}$* embryos, a combination effect could lead to abnormal *Nodal* expression, resulting in embryonic lethality. However, analysis of Nodal and *Cyp26a1* in *Gdf11* mutants and *Nodal*, *Cripto* and *T* expression in *Gdf11$^{-/-}$::T-streak-Cre^{ERT}#47$^{+/0}$* and in gastrulating *T-streak-Cre^{ERT}#47$^{+/+}$* embryos by ISH revealed no apparent differences when compared to wild type littermates **(Fig. 26)**, indicating that neither altered RA signalling

nor abnormal Nodal expression could justify the apparent early lethality of *Gdf11*$^{-/-}$*::T-streak-Cre*^{ERT}*#47*$^{+/0}$ embryos.

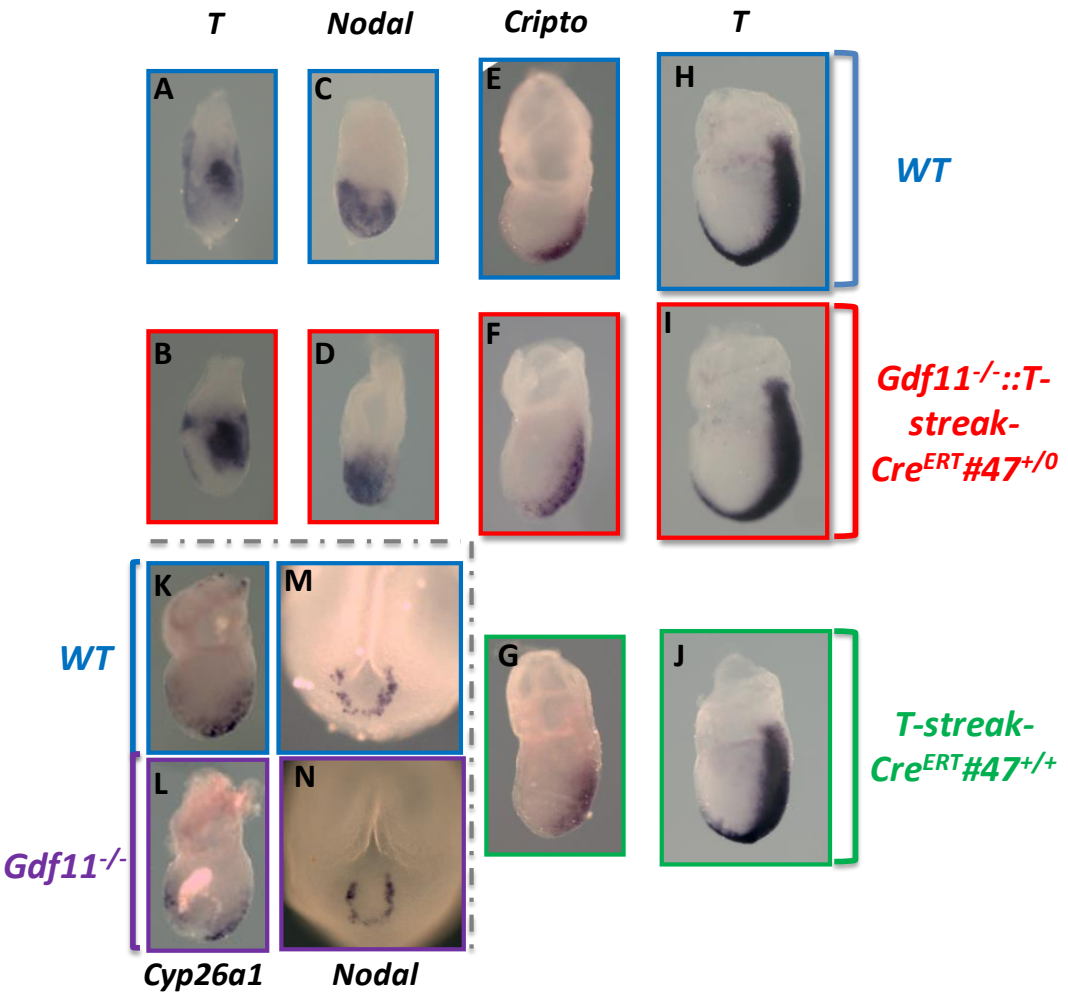


Fig. 26 – No significant differences were found in comparative gene expression analysis considering *Gdf11*$^{-/-}$, *Gdf11*$^{-/-}$*::T-streak-Cre*^{ERT}*#47*$^{+/0}$, *T-streak-Cre*^{ERT}*#47*$^{+/+}$ and wild type embryos during gastrulation.

Although these are just preliminary results, they clearly indicate that at least some *Gdf11*$^{-/-}$*::T-streak-Cre*^{ERT}*#47*$^{+/0}$ embryos survived through gastrulation and formed a defined PS as well as the three germ layers. With the exception of a possible misregulation of *Pcdh8* in *#47* embryos at E7.5, expression of all other genes analysed was indistinguishable between *Gdf11*$^{-/-}$*::T-streak-Cre*^{ERT}*#47*$^{+/0}$ embryos, *Gdf11*$^{-/-}$, *T-streak-Cre*^{ERT}*#47*$^{+/+}$ and wild type littermates. These results were surprising because not only did they fail to suggest any combinatorial effect between the *Gdf11* mutation and the *T-streak-Cre*^{ERT} transgene in the *#47* line, but also did not confirm the apparent early lethality of *Gdf11*$^{-/-}$*::T-streakCre*^{ERT}*#47*$^{+/0}$, since these embryos were found at developmental stages at which earlier results suggested that they were already dead.

To further check this observation, *Gdf11*$^{+/-}$*::T-streak-Cre*^{ERT}*#47*$^{+/+}$ females were again crossed with *Gdf11*$^{+/-}$ males and this time the embryos were harvested at E8.5. Surprisingly, we found *Gdf11*$^{-/-}$*::T-streak-Cre*^{ERT}*#47*$^{+/0}$ embryos within these litters, with apparently normal phenotypes. The same happened at E9.5. Complementary crosses were also performed (*Gdf11*$^{+/-}$ females x *Gdf11*$^{+/-}$*::T-streak-Cre*^{ERT}*#47*$^{+/+}$ males) to rule out the existence of some type of sex-linkage in the lethal phenotype. Normal *Gdf11*$^{-/-}$*::T-streak-Cre*^{ERT}*#47*$^{+/0}$ embryos could be also obtained from such crosses at E9.5 and E10.5 **(Fig. 27)**, which had phenotypic characteristics similar to those observed in *Gdf11* mutant embryos. Also, although the sample size is still small, distribution of different genotypes in *Gdf11::T-streak-Cre*^{ERT}*#47* embryos are compatible with the expected Mendelian ratios **(Table 11)**.

Fig. 27 - *Gdf11*$^{-/-}$*::T-streak-Cre*^{ERT}*#47*$^{+/0}$ embryo at E10.5

Genotype/Stage	E6.5/E7.0	E7.5	E8.0/E8.5	E9.5	E10.5	Total (102 embryos)	Frequency (%)	
							Observed	Expected
Gdf11+/+	4	6	7	2	2	21	20,59	25
Gdf11+/-	9	22	17	5	5	58	56,86	50
Gdf11-/-	5	7	6	4	1	23	22,55	25

The fit to Mendelian expectations was tested with chi-square test: χ^2=0,37, degrees of freedom =2, p>0,9.

Table 11 – *Gdf11* genotyping in the *T-streak-Cre*^{ERT}*#47* background

Together these studies show that the *T-streak-Cre*^{ERT}*#47* transgene has no negative effects on embryonic development either in a wild type or in a *Gdf11* mutant background thus

failing to confirm the previously described early lethality of *Gdf11*$^{-/-}$*::T-streak-Cre*^{ERT}*#47*$^{+/0}$ embryos.

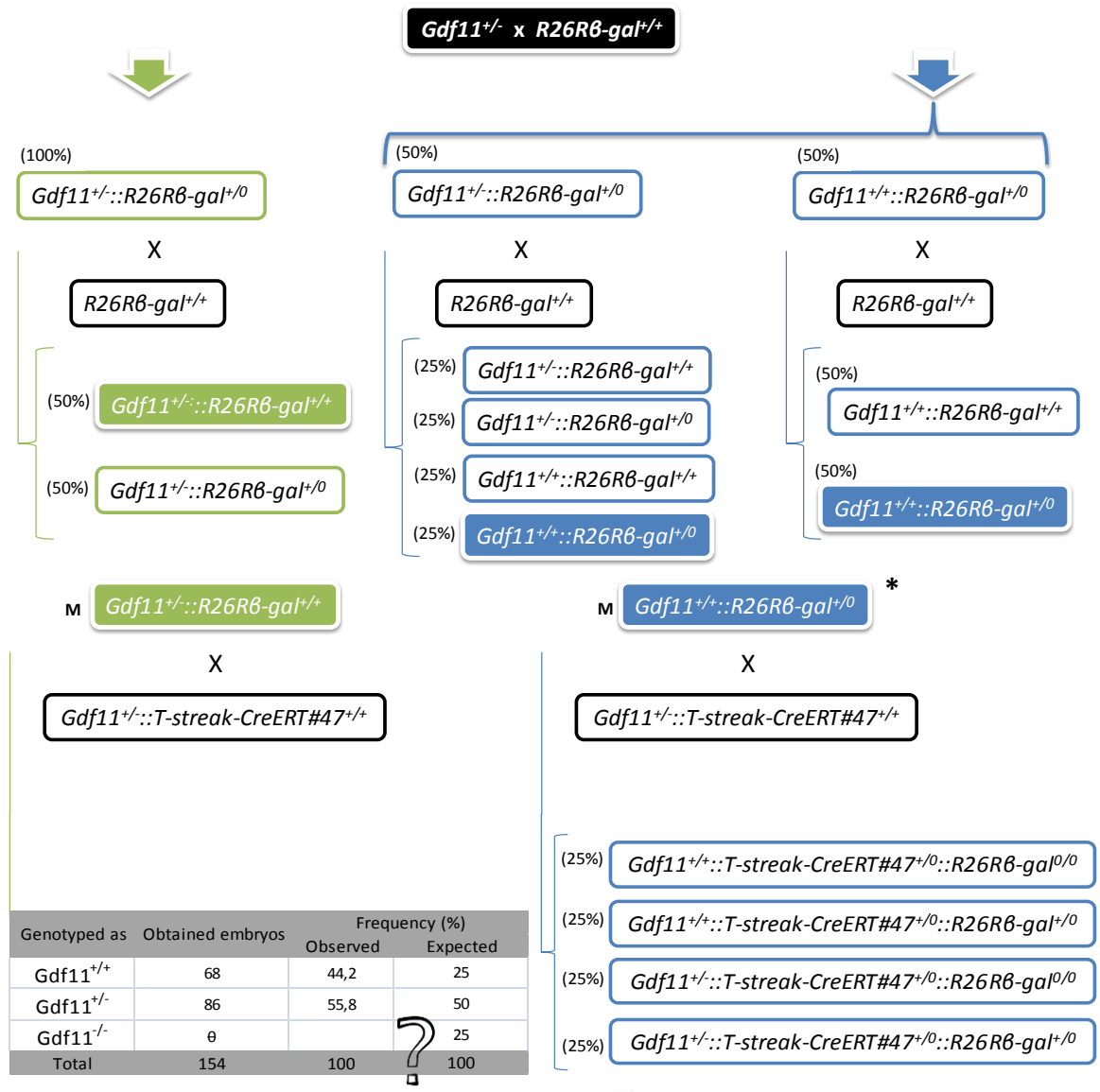
Discussion

Gdf11$^{-/-}$T-streak-Cre^{ERT}#47$^{+/0}$ at E10.5 - a twist of fate

In this work we tried to identify the causes for the early lethality of *Gdf11$^{-/-}$::T-streak-Cre^{ERT}#47$^{+/0}$* embryos. Surprisingly, the results not only did not allow the identification of a cause for such lethality but actually failed to confirm the previous observations with these embryos. The only alteration observed in embryos carrying the *T-streak-Cre^{ERT}#47* transgene was a possible small difference in early *Pcdh8* expression. However, this apparent alteration seemed to have no negative effect on embryonic development, since no combinatorial effect was detectable between the absence of *Gdf11* signalling and the *T-streak-Cre^{ERT}#47* transgene. Consistent with this, *Gdf11$^{-/-}$::T-streak-Cre^{ERT}#47$^{+/0}$* embryos were found to be viable at E10.5 and following the expected Mendelian distribution. Indeed, we could recover living *Gdf11$^{-/-}$::T-streak-Cre^{ERT}#47$^{+/0}$* embryos at this stage regardless of whether the transgene was provided by the father or the mother, therefore ruling out a sex-specific effect. Nevertheless it is still possible that a small proportion of *Gdf11$^{-/-}$::T-streak-Cre^{ERT}#47$^{+/0}$* embryos died early in development but if this was the case, the respective proportion should be very small.

The neo cassette theory

What is then the reason for the differences between the observations of previous and the present work? One possible explanation for the complete absence of *Gdf11$^{-/-}$::R26R-ß-gal$^{+/0}$::T-streakCre^{ERT}#47$^{+/0}$* at E10.5 was that the primers used for genotyping the *Gdf11* mutant allele in the parental *Gdf11$^{+/-}$::ROSA26R-ß-gal$^{+/0}$* lines, were not unique for this allele. We have recently realized that the primers used to identify the mutant *Gdf11* allele, designed to amplify a part of the neomycin cassette inserted to inactivate the *Gdf11* gene, also recognize a similar sequence in the *ß-gal* portion of the *ROSA26R-ß-gal* reporter. This means that when mice were genotyped to identify those with the *Gdf11::ROSA26R-ß-gal* genotype, Jurberg could have mistakenly considered some *Gdf11$^{+/+}$::ROSA26R-ß-gal$^{+/0}$* mice as *Gdf11$^{+/-}$::ROSA26R-ß-gal$^{+/0}$*. In this case, if *Gdf11$^{+/+}$::ROSA26R-ß-gal$^{+/0}$* males were used instead of *Gdf11$^{+/-}$::ROSAROSA26R-ß-gal$^{+/+}$* to cross with *Gdf11$^{+/-}$::T-streak-Cre^{ERT}#47$^{+/0}$* females, it would be impossible to retrieve *Gdf11$^{-/-}$::ROSA26R-ß-gal$^{+/0}$::T-streakCre^{ERT}#47$^{+/0}$* embryos at any stage **(Fig. 28)**. As the wrongly genotyped parental strain could provide the males used in the experiments by Arnon Jurberg, the same wrongly genotyped mice could had been repeatedly used in the different experiments, thus reproducing the mistake in each experiment. Still, it is possible (although highly unlikely) that it was the introduction of the *R26R-ß-gal* background what contributed to the lethality of the embryos in the original experiments. Therefore, to formally prove that in fact *Gdf11$^{-/-}$::ROSA26R-ß-gal$^{+/0}$::T-streakCre^{ERT}#47$^{+/0}$* embryos are viable at E10.5, a new set of primers was designed to detect specifically the *Gdf11* mutant allele. Current efforts in the lab are being made to repeat the initial experiments using this new genotyping strategy.



Genotyped as	Obtained embryos	Frequency (%) Observed	Expected
Gdf11$^{+/+}$	68	44,2	25
Gdf11$^{+/-}$	86	55,8	50
Gdf11$^{-/-}$	0		25
Total	154	100	100

Genotyped as	Obtained embryos	Frequency (%) Observed	Expected*	
Gdf11$^{+/+}$	68	44,2	~~25~~	25
Gdf11$^{+/-}$	86	55,8	~~50~~	75
~~Gdf11$^{-/-}$~~	~~0~~		~~25~~	
Total	154	100	100	

The fit to Mendelian expectations was tested with chi-square test: $\chi 2$=0,000009248, degrees of freedom =1, p>0,95.

Fig. 28 – Mouse mating scheme as designed by Arnon Jurberg (green) and the results of the actual experiments (blue). Considering that he had chosen a progenitor male with only one reporter allele (*), this hypothesis fits Mendelian ratios with a probability higher than 95%.

Acknowledgments/Contributions

André Dias performed most of the experiments reported in this chapter with some contributions from members of the Mallo lab and collaborators. In particular, Rita Aires helped in several experimental procedures and contributed to the research design.

Chapter IV

Final Remarks

Final remarks

The general aim of this MSc thesis was to better understand how RA regulates vertebrate embryonic development. The work presented here provided evidence that RA is only necessary during the initial phase of axial elongation. RA activity in the axial progenitor cells seems to set the time at which HTT occurs. The controlling mechanism of such transition is still unknown, but could be related to Wnt signalling. After the HTT, Cyp26a1 protects NMPs from RA activity, allowing the embryo to continue its growth. In the case of Gdf11 mutants, due to abnormalities in the referred protection, an excess of RA in the tailbud results in an incorrect trunk to tail transition. However, although the RA-Cyp26a1 balance is changed in the absence of Gdf11 signalling at this developmental stage, according to the analysis described in this work, it does not change this balance during gastrulation, thus reinforcing the fact that RA acts in a tissue, time and dose-specific manner. Future studies will be necessary to better understand the role of RA during the HHT and how alterations in this signalling could lead to congenital diseases. It would be interesting to find out if these understanding could lead to possible therapeutic approaches to those diseases using external modulation of RA signalling levels during pregnancy.

Agradecimentos

Em primeiro lugar, gostaria de agradecer ao Moisés por me ter aceitado no seu laboratório e principalmente pela forma como me permitiu desenvolver estes dois projetos de investigação. Tenho ainda a agradecer-lhe a honestidade, a confiança, a preocupação, o facto de ter estado sempre presente, os conhecimentos teóricos e práticos que me transmitiu e ajuda que me deu na escrita desta tese.

A todos os membros do laboratório, em especial à Rita, gostaria de agradecer pela preocupação, pela entreajuda, por todos os ensinamentos que me transmitiram e principalmente por me terem ajudado a perceber um pouco melhor o que é isto de ser investigador, fazer ciência e lidar com o que daí advém.

As palavras de apoio e incentivo, a disponibilidade e os comentários durante escrita da tese, feitos pela Professora Solveig, foram também importantes durante este 2º ano do Mestrado em BED.

Por último, gostaria de agradecer a todos aqueles que de uma forma ou de outra, indiretamente me ajudaram a chegar até este ponto, em especial aos meus pais por tudo o que fizeram por mim.

References

1. Retinoic Acid Signaling. Cell 147, 2011–2013 (2011).

2. Retinoic acid in development: towards an integrated view. Nat.Rev.Genet. 9, 541–53 (2008).

3. Retinoic acid signalling during development. Development 139, 843–58 (2012).

4. Vitamin A deficiency induces congenital spinal deformities in rats. PLoS One 7, e46565 (2012).

5. Retinoic acid-induced spina bifida: evidence for a pathogenetic mechanism. Development 108, 73–81 (1990).

6. Retinoic Acid: Regulation of the Somitogenesis Clock. Birth Defects Res. C. Embryo Today 2153, 205–247 (2008).

7. Removal of maternal retinoic acid by embryonic CYP26 is required for correct Nodal expression during early embryonic patterning. Genes Dev. 23, 1689–98 (2009).

8. Extension and patterning of the vertebrate body. Arnon Jurberg PhD Thesis, Mallo lab (2013).

9. Regulation of anterior/posterior patterning of the axial skeleton by growth differentiation factor 11. Nature 22, 1–5 (1999).

10. Switching axial progenitors from producing trunk to tail tissues in vertebrate embryos. Dev. Cell 25, 451–62 (2013).

11. Cdx2 is essential for axial elongation in mouse development. Proc. Natl. Acad. Sci. USA 101, 7641–7645 (2004).

12. Nadine Dobrovolskaia-Zavadskaia and the dawn of developmental genetics. BioEssays 23, 365–371 (2001).

13. Cdx mediates neural tube closure through transcriptional regulation of the planar cell polarity gene Ptk7. Development 138, 1361–1370 (2011).

14. Evolutionarily conserved requirement of Cdx for post-occipital tissue emergence. Development 139, 2576–2583 (2012).

15. Homeosis and intestinal tumours in Cdx2 mutant mice. Nature 386, 84–87 (1997).

16. Brachyury--a gene affecting mouse gastrulation and early organogenesis. Dev. Suppl. 165, 157–165 (1992).

17. Wnt-3a regulates somites and tailbud formation in the mouse embryo. Genes Dev. 8, 174–189 (1994).

18. Mechanisms of retinoic acid signalling and its roles in organ and limb development. Nat. Rev. Molecular Cell Biology 16, 110–123 (2015).

19. Embryonic retinoic acid synthesis is essential for early mouse post- implantation development letter. Nat. Genet. 21, 444-448 (1999).

20. Novel retinoic acid generating activities in the neural tube and heart identified by conditional rescue of Raldh2 null mutant mice. Development 129, 2271–82 (2002).

21. Knockout mouse models to study Wnt signal transduction. Trends Genet. 22, 678–89 (2006).

22. Effect of retinoic acid signaling on Wnt/beta-catenin and FGF signaling during body axis extension. Gene Expr. Patterns 9, 430–435 (2009).

23. The retinoic acid-metabolizing enzyme, CYP26A1, is essential for normal hindbrain patterning, vertebral identity, and development of posterior structures. Genes Dev. 15, 226–240 (2001).

24. 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 (2001).

25. Growth differentiation factor 11 signaling controls retinoic acid activity for axial vertebral development. Dev. Biol. 347, 195–203 (2010).

26. Mouse gastrulation: the formation of a mammalian body plan. Mech. Dev. 68, 3–25 (1997).

27. Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. Nat. Rev. Mol. Cell Biol. 10, 91–103 (2009).

28. Cell fate decisions and axis determination in the early mouse embryo. Development 139, 3–14 (2012).

29. Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. Development 136, 701–713 (2009).

30. Gene function in mouse embryogenesis : get set for gastrulation. Nat. Rev. Genet. 8, 368–381 (2007).

31. Building the mouse gastrula: signals, asymmetry and lineages. Curr. Opin. Genet. Dev. 16, 419–425 (2006).

32. Beta-catenin regulates Cripto- and Wnt3-dependent gene expression programs in mouse axis and mesoderm formation. Development 130, 6283–94 (2003).

33. Nodal Expression in the Primitive Endoderm Is Required for Specification of the Anterior Axis During Mouse Gastrulation. Development 124, 1033–44 (1997).

34. A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. Development 120, 1919–28 (1994).

35. Wnt3 signaling in the epiblast is required for proper orientation of the anteroposterior axis. Dev. Biol. 312, 312–20 (2007).

36. The Wnt co-receptors Lrp5 and Lrp6 are essential for gastrulation in mice. Development 131, 2803–2815 (2004).

37. Requirement for Wnt3 in vertebrate axis formation. Nat. Genet. 22, 361–365 (1999).

38. Wnt3 function in the epiblast is required for the maintenance but not the initiation of gastrulation in mice. Dev. Biol. 130, 9492–9499 (2013).

39. Requirement for beta-catenin in anterior-posterior axis formation in mice. J. Cell Biol. 148, 567–78 (2000).

40. SnapShot: mouse primitive streak. Cell 146, 488–488.e2 (2011).

41. Targeted disruption of Fgf8 causes failure of cell migration in the gastrulating mouse embryo. Genes Dev. 13, 1834–1846 (1999).

42. Regionally specific induction by the Spemann-Mangold organizer. Nat. Rev. Genet. 5, 425–34 (2004).

43. The Key to Left-Right Asymmetry. Cell 127, 27–32 (2006).

44. The left-right axis in the mouse: from origin to morphology. Development 133, 2095–2104 (2006).

45. Left–right asymmetry in embryonic development: a comprehensive review. Mech. Dev. 122, 3–25 (2005).

46. Differences in left-right axis pathways in mouse and chick: functions of FGF8 and SHH. Science 285, 403–6 (1999).

47. Nodal activity in the node governs left-right asymmetry. Genes&Development 16, 2339–2344 (2002).

48. Wasiak, S. & Lohnes, D. Retinoic acid affects left-right patterning. Dev. Biol. 215, 332–42 (1999).

49. Retinoic acid is required in the mouse embryo for left-right asymmetry determination and heart morphogenesis. Development 126, 2589–96 (1999).

50. Retinoic-acid signalling in node ectoderm and posterior neural plate directs left – right patterning of somitic mesoderm. Nat. Cell Biol. 8, 271–7 (2006).

51. Coupling segmentation to axis formation. Development 131, 5783–93 (2004).

52. Deciphering the function of canonical Wnt signals in development and disease: conditional loss- and gain-of-function mutations of beta-catenin in mice. Genes Dev. 22, 2308–41 (2008).

53. Wnt3a/β-catenin signaling controls posterior body development by coordinating mesoderm formation and segmentation. Development 135, 85–94 (2007).

54. Regulation of canonical Wnt signaling by Brachyury is essential for posterior mesoderm formation. Dev. Cell 15, 121–33 (2008).

55. T (Brachyury) is a direct target of Wnt3a during paraxial mesoderm specification. Genes Dev. 13, 3185–3190 (1999).

56. Signaling by FGF4 and FGF8 is required for axial elongation of the mouse embryo. Dev. Biol. 371, 235–245 (2012).

57. Axial progenitors with extensive potency are localised to the mouse chordoneural hinge. Development 129, 4855–4866 (2002).

58. Neuromesodermal progenitors and the making of the spinal cord. Development 142, 2864–2875 (2015).

59. Two distinct sources for a population of maturing axial progenitors. Development 134, 2829–2840 (2007).

60. Tbx6-dependent Sox2 regulation determines neural or mesodermal fate in axial stem cells. Nature 470, 394–398 (2011).

61. Three neural tubes in mouse embryos with mutations in the T-box gene Tbx6. Nature 391, 695–697 (1998).

62. Region-specific regulation of posterior axial elongation during vertebrate embryogenesis. Dev. Dyn. 243, 88–98 (2014).

63. Loss of FGF-Dependent Mesoderm Identity and Rise of Endogenous Retinoid Signalling Determine Cessation of Body Axis Elongation. PLoS Biol. 10, e1001415 (2012).

64. Lineage tracing of neuromesodermal progenitors reveals novel Wnt-dependent roles in trunk progenitor cell maintenance and differentiation. Development 142, 1628–1638 (2015).

65. Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev. 17, 126–140 (2003).

66. Cdx and Hox Genes Differentially Regulate Posterior Axial Growth in Mammalian Embryos. Dev. Cell 17, 516–526 (2009).

67. Developing with lethal RA levels: genetic ablation of Rarg can restore the viability of mice lacking Cyp26a1. Development 130, 1449–1459 (2003).

68. Respecification of vertebral identities by retinoic acid. Development 115, 487–501 (1992).

69. Segmental patterning of the vertebrate embryonic axis. Nat. Rev. Genet. 9, 370–382 (2008).

70. Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. Cell 91, 639–648 (1997).

71. Segmentation in vertebrates: Clock and gradient finally joined. Genes Dev. 18, 2060–2067 (2004).

72. A beta-catenin gradient links the clock and wavefront systems in mouse embryo segmentation. Nat. Cell Biol. 10, 186–193 (2008).

73. Modulating Hox gene functions during animal body patterning. Nat. Rev. Genet. 6, 893–904 (2005).

74. Shifting boundaries of retinoic acid activity control hindbrain segmental gene expression. Development 5, 2611–2622 (2005).

75. Developmental functions of mammalian Hox genes. Mol. Hum. Reprod. 3, 115–31 (1997).

76. Retinoic Acid Promotes Limb Induction through Effects on Body Axis Extension but Is Unnecessary for Limb Patterning. Curr. Biol. 2, 1050–1057 (2009).

77. RA Acts in a Coherent Feed-Forward Mechanism with Tbx5 to Control Limb Bud Induction and Initiation. Cell Rep. 12, 879–891 (2015).

78. Embryonic retinoic acid synthesis is required for forelimb growth and anteroposterior patterning in the mouse. Development 129, 3563–74 (2002).

79. Retinoic acid coordinates somitogenesis and left – right patterning in vertebrate embryos. Nature 435, 215-220 (2005).

80. Retinoic acid controls the bilateral symmetry of somite formation in the mouse embryo. Science 308, 563–566 (2005).

81. Rere controls retinoic acid signalling and somite bilateral symmetry. Nature 463, 953–957 (2010).

82. Retinoic acid controls heart anteroposterior patterning by down-regulating Isl1 through the Fgf8 pathway. Dev. Dyn. 237, 1627–35 (2008).

83. Opposing FGF and retinoid pathways control ventral neural pattern, neuronal differentiation, and segmentation during body axis extension. Neuron 40, 65–79 (2003).

84. Retinoic acid controls body axis extension by directly repressing Fgf8 transcription. Development 141, 2972–7 (2014).

85. Developmental regulation of the Hox genes during axial morphogenesis in the mouse. Development 132, 2931–42 (2005).

86. Tbx5 is essential for forelimb bud initiation following patterning of the limb field in the mouse embryo. Development 130, 623–633 (2003).

87. Unpublished data from Rita Aires, PhD student in Mallo Lab (2015).

88. Effect of retinoic acid signaling on Wnt/β-catenin and FGF signaling during body axis extension. Gene Expr Patterns 9, 430–435 (2009).

89. Antagonism between Retinoic Acid and Fibroblast Growth Factor Signaling during Limb Development. Cell Rep. 3, 1503–1511 (2014).

90. Early mouse caudal development relies on crosstalk between retinoic acid , Shh and Fgf signalling pathways. Development 676, 665–676 (2009).

91. Transcriptomic Analysis of Murine Embryos Lacking Endogenous Retinoic Acid Signaling. PLoS One 8, e62274 (2013).

92. MesP1 is expressed in the heart precursor cells and required for the formation of a single heart tube. Development 126, 3437–3447 (1999).

93. Genetic rescue of segmentation defect in MesP2-deficient mice by MesP1 gene replacement. Mech. Dev. 75, 53–66 (1998).

94. Mesp1: A Key Regulator of Cardiovascular Lineage Commitment. Circ. Res. 107, 1414–1427 (2010).

95. Dkk1 and Wnt3 interact to control head morphogenesis in the mouse. Development 135, 1791–1801 (2008).

96. Dickkopf1 Is Required for Embryonic Head Induction and Limb Morphogenesis in the Mouse. Dev. Cell 1, 423–434 (2001).

97. Expression pattern of Dkk-1 during mouse limb development. Mech. Dev. 89, 151–3 (1999).

98. Genetic interaction of Gsc and Dkk1 in head morphogenesis of the mouse. Mech. Dev. 124, 157–165 (2007).

99. Dickkopf genes are co-ordinately expressed in mesodermal lineages. Mech. Dev. 87, 45–56 (1999).

100. A complex oscillating network of signaling genes underlies the mouse segmentation clock. Science 314, 1595–1598 (2006).

101. Caveolin-1 expression inhibits Wnt/beta-catenin/Lef-1 signaling by recruiting beta-catenin to caveolae membrane domains. J. Biol. Chem. 275, 23368–23377 (2000).

102. Caveolin-1 Null Mice Are Viable but Show Evidence of Hyperproliferative and Vascular Abnormalities. J. Biol. Chem. 276, 38121–38138 (2001).

103. Tissue-specific roles of Axin2 in the inhibition and activation of Wnt signaling in the mouse embryo. Proc. Natl. Acad. Sci. 108, 8692–8697 (2011).

104. Expression of growth/differentiation factor 11, a new member of the BMP/TGFbeta superfamily during mouse embryogenesis. Mech. Dev. 80, 185–9 (1999).

105. Mouse paraxial protocadherin is expressed in trunk mesoderm and is not essential for mouse development. Genesis 27, 49–57 (2000).

106. Xenopus paraxial protocadherin inhibits Wnt/β-catenin signalling via casein kinase 2β. EMBO Rep. 13, 129–34 (2012).

List of Abbreviations

AP	anterior-posterior
AVE	anterior visceral endoderm
Bmp	bone morphogenetic protein
Bhlhe40	basic helix-loop-helix family, member e40
Cav1	caveolin 1
Cdx	caudal-type homeobox
Cerl1	cerberus 1 homolog (Xenopus laevis)
Cyp26a1	cytochrome P450, family 26, subfamily a, polypeptide 1
CNH	chordoneural hinge
Crabp2	cellular retinoic acid binding protein II
Cre^{ERT}	tamoxifen-inducible Cre recombinase
Dkk1	dickkopf homolog 1 (Xenopus laevis)
DNA	deoxyribonucleic acid
Dusp6/Mkp3	dual specificity phosphatase 6
DVE	Distal visceral endoderm
e.g.	*exempli gratia*
Eno2	enolase 2
Fgf	fibroblast growth factor
Gata	GATA binding protein
Gdf	growth and differentiation factor
Hand	heart and neural crest derivatives
Hox	homebox
HTT	head to trunk transition
ICM	Inner cell mass
ISH	in situ hybridization
Isl1	islet1
Lef1	lymphoid enhancer binding factor 1
Lefty1	left right determination factor 1
Lfng	lunatic fringe
MABT	maleic acid buffer containing Tween 20
Meis	meis homebox
Meox	mesenchyme homeobox
Mesp	mesoderm posterior
Nkx1-2	NK1 transcription factor related, locus 2 (Drosophila)
NMPs	neural mesodermal progenitors
NSB	node streak border
Olfm4	olfactomedin 4
ON	over-night
Oct4/Pou5f1	octamer-binding transcription factor 4/POU domain class 5, transcription factor 1
Pax	paired box
PBS	phosphate buffered saline
PBT	phosphate buffered saline with Tween-20
Pcdh8/Papc	protocadherin 8 / paraxial protochaderin

PCR	polymerase chain reaction
PFA	paraformaldehyde
PS	primitive streak
PSM	presomitic mesoderm
RA	retinoic acid
Raldh2/Aldh1a2	aldehyde dehydrogenase family 1, subfamily A2
RNA	ribonucleic acid
Shh	sonic hedgehog
Sox 2	SRY-box 2
Spry2	sprouty homolog 2 (Drosophila)
T	brachyury
TE	trophectoderm
TTT	trunk to tail transition
Tbx6	t-box
Tcf1	transcription factor 1, T cell specific
Wif1	Wnt inhibitory factor 1
Wnt	wingless-type MMTV integration site family
X-gal	5-bromo-4-chloro-indolyl-β-D-galactopyranoside