

UNIVERSIDADE DE LISBOA  
FACULDADE DE CIÊNCIAS  
DEPARTAMENTO DE BIOLOGIA ANIMAL



## **Functional analysis of candidate genes affecting Hoxa10 activity**

André Daniel Faustino Mesquita

**Mestrado em Biologia Evolutiva e do Desenvolvimento**

Dissertação orientada por:  
Doutor Moises Mallo Perez  
Professora Doutora Sólveig Thorsteinsdóttir



## Acknowledgments

O primeiro agradecimento tem de ir para o Moises. Deu-me uma oportunidade. Às cegas, aceitou-me sem me conhecer de lado nenhum e nunca me colocou barreiras. Com a porta sempre aberta para tirar dúvidas e dar sugestões. Acima de tudo agradeço pela maneira como sempre foi honesto comigo, dizendo aquilo que precisei de ouvir no momento certo.

Depois vem a Professora Sólveig, que aceitou ser minha orientadora interna. Esteve sempre presente quando precisei de ajuda e principalmente porque é uma excelente ouvinte. Além de professora, foi quase como uma psicóloga.

A seguir é a vez do pessoal do laboratório que foi uma grande ajuda. Ana Casaca: foi quem passou mais tempo comigo e me ensinou muitas coisas. Não tanto pelas técnicas, mas sim pela maneira como pensar, a ser mais observador, mais rigoroso. Acima de tudo ensinaste-me a ser paciente. Rita: sempre que tive dúvidas, nunca me deixaste mal. Ensinaste-me a fazer umas belas imunos. Se saíram belas ou não, isso é que já não sei. Ah, mostraste-me essa pérola chamada “Ninja das Caldas”, algo que nunca vou conseguir apagar da minha memória. Luísa: ensinaste-me muitas coisas, mas o que mais te agradeço é teres sido mais do que uma colega. Senti que querias genuinamente saber como eu e o meu trabalho estávamos. Não viste um único filme sugerido por mim, mas eu vou deixar essa passar... ainda vais a tempo! Ana Nóvoa: aqueles transgénicos não se microinjectam sozinhos, graças a essas mãos de ouro tenho uma tese para apresentar! Além desta parte óbvia, admiro a tua personalidade. Frontal, directa, sem medo de dizeres o que pensas. Fazem falta pessoas assim. Irma: poucas pessoas são tão fortes como tu. No pouco tempo que partilhámos no laboratório foi fácil perceber que és um exemplo como pessoa. Agradeço também pelos Western Blots que me ensinaste a fazer. Logo aí foi metade da minha tese. André: não foi muito o tempo juntos no laboratório, mas mesmo assim ajudaste. Um grande obrigado por me “orientares” antes de aí chegar e sempre que precisei. Tiago, és o membro honorário do Mallo lab e é bem merecido. Foste um grande conselheiro e agradeço-te por toda a ajuda que deste.

Porque a minha vida não foi só IGC, tenho de agradecer aqueles que foram, são e continuarão a ser parte da minha vida. Inês, és uma grande amiga. A tua paciência para ouvir os meus desabafos é invejável. A minha sanidade mental deve-se a ti! Rita, este ano pouco falámos, mas desde que te conheci naquela aula de Citogenética tens sido uma grande amiga. Teresinha, só nos conhecemos no mestrado e deu para ver logo que és diferente (in a good way?), quando for grande quero ser como tu. Rodrigão! És tão boa pessoa que faz impressão. És um amigo, um espectáculo e acredito mesmo que vais ser um grande cientista no futuro! Zé... Já lá vão uns bons anos. Nem me lembro do momento em que nos conhecemos, mas ainda bem que aconteceu. Ano após ano, cá estamos. Tens uma paciência incrível para me aturares, espero que não te fartes!

David, Daniel, Catarina e Alberto, a vossa presença é essencial. Tenho um orgulho enorme de ano após ano continuar a crescer com vocês ao meu lado, viver coisas novas e divertir-me como nunca.

Ciro e Albuquerque, dois ídolos. Dois mentores e grandes amigos. Aprendi e continuo a aprender muito com vocês.

Por fim, o maior agradecimento vai para a minha família. Mãe e Pai, vocês são uma inspiração. Vieram do nada e ainda assim deram-me todas as oportunidades que podia pedir, a mim e ao Bruno (a minha outra grande inspiração). Não tenho como vos agradecer, só vos quero ver orgulhosos e felizes (vou dar o meu melhor!). Nádia... Estás sempre a meu lado, no bom e mau, dizes sempre a coisa certa e motivas-me como ninguém. Espero que esta aventura continue durante muito tempo!



## Abstract

*Hox* genes encode transcription factors that control axial patterning in all bilaterians. They are characterized by the presence of a protein motif, the homeodomain (HD), which is responsible for the physical interaction between Hox proteins and their DNA targets. In vertebrates, *Hox10* genes have rib-repressing activity, which determines the thoracic to lumbar transition, a functional property not shared by any other Hox protein. Previous work showed that Hox10 functional specificity does not reside in their HD but requires input from other parts of the protein. This includes a conserved motif, known as C1. A yeast two-hybrid screen identified several factors potentially interacting with C1, which could be candidates for Hox10 functional cofactors. Here, we analysed some of those factors, including Grg3, IFT144 and Smad4 for their ability to interact with Hoxa10 in a co-immunoprecipitation assay in cultured cells. These experiments failed to detect interactions between any of these proteins and Hoxa10, thus arguing against them being Hox10 cofactors. In addition, we created a new mutant line for Grg3 and analysed its role in skeletal formation. These analyses revealed that the axial skeleton in general and ribs in particular form in the absence of Grg3, thus reinforcing the conclusion that this protein is not a functional cofactor of Hox10 proteins. In addition to this, we further tested the role of the C1 motif for the rib-repressing function of Hox10 proteins by testing several deletion mutants in transgenic mouse embryos. These experiments indicated that C1 plays a role in Hox10 rib-blocking function. In addition, when the C1 motif was introduced into Hoxb9, the chimeric protein blocked rib formation in transgenic embryos, a property absent from the native Hoxb9 protein. These experiments showed that the C1 motif is also sufficient to promote a rib-repressing function. Surprisingly, these embryos also contained skeletal phenotypes consistent with abnormal segmentation of the paraxial mesoderm. These data suggest that the C1 motif might interact with the segmentation clock opening the possibility that regulation of rib formation might occur by modulating specific features of the segmentation network.

**Key words:** *Hox* genes; Grg3; Smad4; IFT144; C1 motif.



## Resumo

Os genes *Hox* codificam factores de transcrição que controlam a padronização axial em todos os Bilatéria. São caracterizados pela presença de um motivo proteico, o homeodomínio (HD), responsável pela interacção física entre proteínas Hox e os seus alvos. Em vertebrados, existem 39 genes *Hox* distribuídos por treze grupos parálogos e organizados em quatro grupos (A, B, C e D). Uma vez que o motivo HD é necessário para a função das proteínas Hox, seria de esperar que também fosse responsável pelas diferentes funções de cada grupo parálogo. No entanto, este não é o caso dado que HDs de diferentes grupos ligam sequências muito semelhantes. De facto, foi relatado que a especificidade funcional está relacionada com regiões fora do HD. Um bom exemplo disso é a proteína Hoxa10, parte do grupo Hox10. Composto por três proteínas, este grupo foi associado ao desenvolvimento correto da região lombar. Sobreexpressar apenas um membro do grupo (Hoxa10) na mesoderme pré-somítica é suficiente para originar embriões de ratinho (*M. musculus*) sem costelas. A inactivação dos três membros produz o resultado oposto, com costelas ao longo da região torácica e lombar. Portanto, este grupo não tem apenas uma actividade de repressão de formação das costelas, mas é também essencial na transição torácica para lombar. Também acontece que as três proteínas Hox10 partilham um motivo idêntico entre elas chamado M1. Este motivo de sete aminoácidos está adjacente ao HD e quando foram utilizadas construções com este motivo ausente, a proteína Hoxa10 praticamente perdeu a sua função repressora. Para complementar, o M1 foi inserido no Hoxb9. Neste caso, os embriões resultantes não exibiram fenótipos anormais ao nível das costelas. Em conjunto, estes resultados indicam claramente que o motivo M1 é necessário, mas não é suficiente para conferir uma função repressora ao Hoxa10. Trabalho do laboratório Mallo também identificou um outro motivo que se pensa que interage com outras proteínas, o C1. Apesar de não ser tão conservado entre os elementos do grupo proteico Hox10, um sistema de duplo híbrido em leveduras identificou várias proteínas a interagir com o C1, podendo essas ser cofactores funcionais do grupo Hox10. Várias destas proteínas continham domínios WD40, uma propriedade estrutural que serve como “molde” para interacções entre proteínas ou entre ADN e proteínas. Também foram identificadas outras proteínas sem estes domínios. Este trabalho teve dois grandes objectivos: 1) analisar três candidatos a cofactores funcionais de Hoxa10 e 2) compreender a importância do motivo C1 na especificidade funcional do grupo Hox10.

A primeira proteína candidata estudada foi a Grg3. Esta foi uma das proteínas detectadas que continham domínios WD40. É uma co-repressora da transcrição e foi reportada a expressão da sua homóloga humana, *TLE3*, no esclerótomo, o precursor das vértebras e costelas. Foram usadas duas abordagens para testar se a Grg3 poderia estar a interagir com o Hoxa10. Primeiro, células humanas (293T) foram transfectadas com os domínios WD40 da Grg3 e também com Hoxa10. Os extractos celulares foram posteriormente usados numa co-imunoprecipitação (Co-IP). Estas experiências não foram capazes de detectar Grg3 no imunoprecipitado. Além disto, utilizando técnicas de imunocitoquímica, confirmámos que a Grg3 estava localizada no núcleo. Em conjunto, estes dados indicam que Hoxa10 e Grg3 não interagem uns com os outros. A segunda abordagem consistiu em eliminar a expressão de *Grg3* usando o sistema CRISPR/Cas9. Os embriões homozigotos E18.5 não revelaram qualquer tipo de alterações no fenótipo das costelas e a expressão de *Myf5*, um marcador muscular induzido por Hoxa10, manteve-se inalterada em embriões E10.5. Logo, estes resultados reforçam a ideia que a Grg3 não interage com Hoxa10.

A IFT144 também foi transfectada em conjunto com Hoxa10. Esta proteína de transporte intraflagelar contém vários domínios WD40 e foi reportado que sua inactivação resulta em defeitos no desenvolvimento das costelas e no início da somitogénese. Ainda que não tenha sido detectada no sistema de duplo híbrido tendo em conta suas características estruturais e funcionais também foi testada. Primeiro, a expressão de *IFT144* foi estudada e, apesar de ser expressa de forma ubíqua, está localizada

na mesoderme pré-somítica. Em segundo lugar, a localização da proteína foi confirmada no núcleo, bem como no citoplasma. Em terceiro lugar, a técnica da Co-IP foi novamente usada e revelou que a IFT144 não parece estar a interagir com Hoxa10.

Smad4, outro possível cofactor do Hoxa10, faz parte da via de sinalização TGF- $\beta$  e é conhecido por fazer parte de complexos que actuam como factores de transcrição. Embora não contenha nenhum domínio WD40, foi uma das proteínas detectadas no sistema de duplo híbrido e também está associada à padronização do eixo antero-posterior. Neste caso foi usada apenas a abordagem da transfecção seguida pela Co-IP. Mais uma vez, não foi detectada qualquer tipo de interação entre as duas proteínas.

Como já foi mencionado anteriormente, o segundo objectivo deste trabalho passava por compreender a importância do motivo C1 na especificidade funcional do grupo Hox10. Para tal, foram utilizadas diferentes construções transgénicas, que continham diferentes versões da proteína Hoxa10. A maior parte dos ratinhos transgénicos que sobreexpressavam uma versão do Hoxa10 que não possuía o motivo C1 (*DllHoxa10 $\Delta$ C1*) não desenvolveram costelas. Alguns destes transgénicos tinham fenótipos menos severos (faltavam algumas costelas), algo que nunca foi reportado em trabalhos onde foi feita a sobreexpressão de Hoxa10. Logo, pode-se concluir que a função de repressão do Hoxa10 foi parcialmente perdida, sugerindo que C1 tem um papel na mesma. Esta ideia foi reforçada por outras duas construções que não possuíam diferentes partes do motivo C1. *DllHoxa10 $\Delta$ C1p2* deu origem a embriões E18.5 com fenótipos pouco severos, enquanto *DllHoxa10 $\Delta$ C1p1* apresentava um embrião sem costelas. O resultado mais interessante veio da construção transgénica *DllHoxb9insC1* quando o motivo C1 foi colocado na proteína Hoxb9, uma proteína que normalmente não tem efeito na formação de costelas. Neste caso foram observados vários embriões sem costelas. Pode-se assim concluir que o C1 não é apenas necessário, mas é também suficiente para conferir uma função repressora da formação de costelas. Mais surpreendente, foi o facto de alguns destes embriões exibirem graves defeitos na segmentação. Esta ideia foi corroborada pelas hibridações *in situ* em embriões E11.5, que praticamente perderam a expressão de *Tbx18* e *Uncx4.1* e com *Paraxis* a evidenciar uma segmentação anormal da mesoderme paraxial. Estes dados sugerem que o motivo C1 poderá interagir com o relógio de segmentação, abrindo a possibilidade que a regulação da formação de costelas pode ocorrer através da modulação de características específicas da rede de segmentação.

**Palavras chave:** *Hox* genes; Grg3; Smad4; IFT144; Motivo C1.







# Table of contents

Acknowledgments.....	III
Abstract.....	V
Resumo.....	VII
Table and Figures Index.....	XIII
I. Introduction.....	1
I.1 The axial skeleton formation .....	1
I.1.1 Somitogenesis: Segmentation.....	1
I.1.2 Somitogenesis: Patterning.....	1
I.1.3 Vertebrae and ribs development .....	1
I.2 Hox genes.....	2
I.2.1 What are <i>Hox</i> genes?.....	2
I.2.2 Vertebrate Hox genes .....	3
I.2.3 The <i>Hox</i> code.....	3
I.2.4 Mouse <i>Hox</i> gene expression .....	3
I.2.5 Hox genes and the axial skeleton development .....	4
I.2.6 Hox protein function and DNA binding properties.....	5
I.3 A molecular look at Hoxa10.....	5
I.3.1 Hoxa10's M1 motif .....	6
I.3.2 Hoxa10's C1 motif .....	6
I.4 Groucho-Related Gene.....	6
I.5 Smad4.....	7
I.6 Objectives .....	8
II. Materials & Methods .....	9
II.1 Animal model.....	9
II.2 Making <i>Grg3</i> knockout mutants .....	9
II.2.1 Genotyping <i>Grg3</i> mutants .....	9
II.3 Making transgenic constructs .....	10
II.3.1 Constructs used in this work.....	10
II.3.2 cDNA synthesis .....	10
II.3.3 Plasmid digestion, isolation and purification for microinjection.....	11
II.4 Microinjection .....	11
II.5 Embryo analysis .....	12
II.5.1 Genotyping.....	12
II.6 Skeletal staining.....	12
II.7 In situ hybridization.....	12
II.7.1 RNA probe synthesis.....	12
II.7.2 In situ procedure .....	13
II.8 Cell culture.....	13
II.8.1 Transfection.....	14
II.8.2 Cell lysis and storage of cell extracts .....	14

II.8.3 Protein extract analysis.....	14
II.9 Co-Immunoprecipitation (Co-IP) .....	14
II.10 Immunostaining.....	15
III. Results.....	17
III.1 Hoxa10 and its functional candidates .....	17
III.1.1 Grg3 .....	17
III.1.1.1 Grg3 Co-IP .....	17
III.1.1.2 Grg3 Knockout mice .....	18
III.1.2 IFT144 .....	19
III.3 Smad4 .....	22
III.2 The C1 motif and the axial skeleton.....	23
IV. Discussion .....	27
References .....	31
Appendix I – Standard Molecular Procedures .....	AI-I
Appendix II: Buffers, Media and Other Solutions .....	AII-I
Appendix III: Sequences and Primers.....	AIII-I
Appendix III: Vector Maps .....	AIV-I

# Table and Figures Index

## Figures:

**Figure 1.1:** *Hox* gene expression and genomic organization in *Drosophila* and *M. musculus*.

**Figure 1.2:** Simplified representation of *Hox* gene expression domains during axial skeleton development.

**Figure 1.3:** Schematic representation of Hox10 proteins and their conserved protein motifs.

**Figure 2.1:** Genotyping strategy of *Grg3* mutants.

**Figure 2.2:** IFT144 cloning strategy.

**Figure 3.1:** Immunostaining of Grg3-WD40 transfected 293T cells.

**Figure 3.2:** Grg3-WD40 Co-IP analysis.

**Figure 3.3:** Alcian blue and alizarin red staining of *Grg3* KO embryos.

**Figure 3.4:** *In situ* hybridization of *Grg3* KO E10.5 embryos.

**Figure 3.5:** Immunostaining of 293T cells for endogenous IFT144.

**Figure 3.6:** *In situ* hybridization showing *IFT144* expression.

**Figure 3.7:** IFT144 Co-IP analysis.

**Figure 3.8:** Smad4 Co-IP analysis

**Figure 3.9:** Schematic representation of each transgenic construct used in the microinjections.

**Figure 3.10:** Representative phenotypes of E18.5 embryos overexpressing different transgenic constructs

**Figure 3.11:** *In situ* hybridizations of E11.5 *DllHoxb9insC1* transgenic embryos with somitogenesis markers.

## Tables:

**Table 2.1:** Proteinase K incubation times for each embryonic stage.



---

# I. Introduction

---

## I.1 The axial skeleton formation

### I.1.1 Somitogenesis: Segmentation

The correct formation of the axial skeleton is a key process in all vertebrates. This critical developmental process begins with the formation of somites in the paraxial mesoderm by a progressive segmentation of the presomitic mesoderm. This happens in a rostral to caudal direction at both sides of the neural tube and under the control of a segmentation clock, which basically consists in a set of genes that are expressed in a cyclic manner, the majority of which are part of the Notch and Wnt pathways (reviewed in Dubrulle and Pourquié, 2004). In addition to this, caudal to rostral gradients of FGF and WNT signalling and, in the opposite direction, of retinoic acid are also involved in somitogenesis (reviewed in Dubrulle and Pourquié, 2004). The combination of the clock and the gradients define the position of the new intersomitic border (reviewed in Dubrulle and Pourquié, 2004).

### I.1.2 Somitogenesis: Patterning

The next step consists in a series of differentiation processes in the newly formed somites, beginning by the dorsal/ventral axis. In their dorsal portion, cells keep an epithelial state and this culminates in the formation of the dermomyotome, which provides precursors for all skeletal muscles (except those of the head), part of the ribs, brown fat cells, some endothelia and dorsal dermis (reviewed in Deries and Thorsteinsdóttir, 2016). In contrast, the ventral portion undergoes an epithelial to mesenchymal transition and forms the sclerotome, which further differentiates into the ventral sclerotome and syndotome (Brent et al., 2003). The first is responsible for the formation of vertebrae, ribs and the intervertebral disks, whilst the second forms the tendons (Brent and Tabin, 2002; Brent et al., 2003).

### I.1.3 Vertebrae and ribs development

As mentioned above, vertebrae arise from the somites, more precisely from the sclerotome. In a first step, the sclerotome undergoes a process known as resegmentation that results in the production of each individual vertebra from the posterior part of one somite and the anterior part of the next. In fact, each vertebra is formed by 2/3 of the posterior ventral portion of the anterior somite and 1/3 of the anterior ventral portion of the posterior somite (Gilbert, 2013).

Mice have seven cervical, thirteen thoracic, six lumbar, four sacral and, usually, twenty-eight caudal vertebrae. This results in a total of fifty-one vertebrae (Burke et al., 1995). The 13 thoracic vertebrae are the ones containing ribs. The sacral vertebrae have rib-like structures, which form the sacrum.

Ribs are classified depending on their connection to the sternum, which originates from the lateral mesoderm (Sudo et al., 2001). The first seven pairs of ribs are considered true ribs (or vertebrosteral ribs) because they are directly attached to the sternum. The eighth to tenth rib pairs are known as false ribs because they are only attached to the sternum by the cartilage (vertebrochondral ribs). The other three pairs are also vertebrochondral, but are referred to as floating ribs because they are only attached to the vertebrae (Srour et al., 2015).

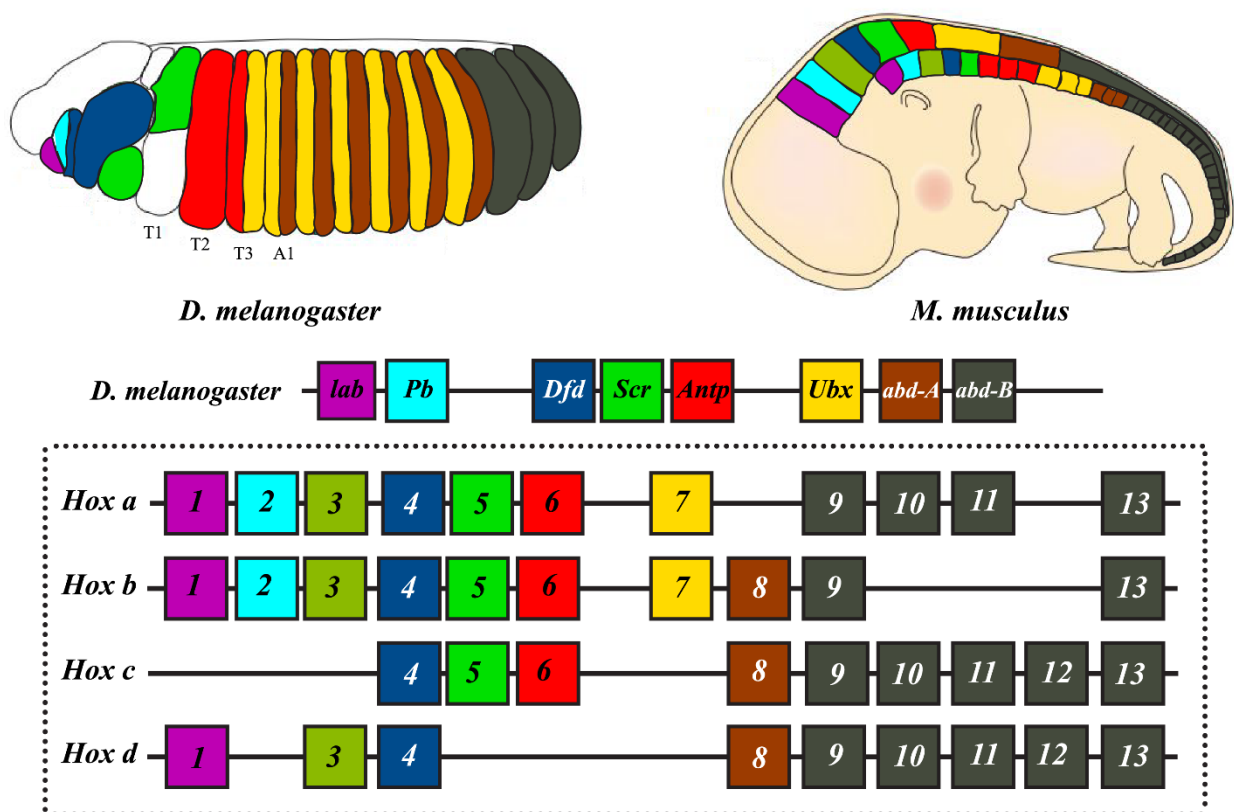
Correct ribcage development is strongly dependent on *Hox* genes. What these genes are and their importance will be addressed below.

## I.2 *Hox* genes

### I.2.1 What are *Hox* genes?

*Hox* genes, or homeotic genes, encode transcriptional regulatory proteins that control axial patterning in all bilaterians (Garcia-Fernández, 2005). They were first discovered in the mid-20<sup>th</sup> century by Edward B. Lewis with his work in *Drosophila melanogaster*. Their name comes from the fact that when mutated they cause homeotic transformations, the transformation of a structure into the likeness of another one. One example of this is the *ANTENNAPEDIA* (*Antp*) mutation, in which the fly develops a pair of legs in the position of the antenna (Lewis, 1978).

In *Drosophila*, besides *Antp*, there are seven other homeotic genes and together they control the identity of each segment, along the anterior/posterior axis (Fig. 1.1). These are: *labial* (*lab*), *proboscipedia* (*Pb*), *Deformed* (*Dfd*), *Sex combs reduced* (*Scr*), *Ultrabithorax* (*Ubx*), *Abdominal-A* (*abd-A*) and *Abdominal-B* (*abd-B*) (reviewed in McGinnis and Krumlauf, 1992). All of them share a highly-conserved DNA sequence known as homeobox (reviewed in McGinnis and Krumlauf, 1992). This 180-base pair (bp) sequence encodes the homeodomain, a protein motif composed by sixty amino acids. It is through this motif that *Hox* proteins interact with their DNA targets and regulate expression of other genes, hence their role as transcription factors (Pearson et al., 2005).



**Figure 1.1: *Hox* gene expression and genomic organization in *Drosophila* and *M. musculus*.** Schematic representation of *Hox* genes expression in *Drosophila* and mouse embryos (top). Thoracic segments (T1–T3) and the first abdominal segment (A1) are labelled in the *Drosophila* embryo. *Hox* gene organization along the chromosome in *Drosophila* is represented below. The four mammalian *Hox* gene clusters are shown inside the dotted box, with each paralog group represented in different colours. Adapted from Pearson et al., 2005.



### **I.2.2 Vertebrate *Hox* genes**

In mammals, there are thirty-nine of these genes organized in four clusters that originated from two whole-genome duplication events (Duboule, 2007). These four clusters are named *Hox A*, *B*, *C* and *D* and their genes are divided into thirteen paralog groups according to sequence similarity and position within the cluster (Fig. 1.1) (Pearson et al., 2005; Wellik, 2007). This means that *Hox* genes of a certain cluster have closer “companions” in other clusters. Other vertebrates have a variable number of clusters. For example, the Zebrafish (*Danio rerio*) instead of four, has seven clusters supposedly from three whole-genome duplications and the subsequent loss of one of those clusters (Woltering and Durston, 2006).

*Hox* genes belonging to the same paralogous group have functional redundancy (Wellik and Capecchi, 2003). This means that in loss-of-function studies in mice it becomes very important to silence all members of a certain group to have a strong enough mutant phenotype. However, when using an overexpression approach only one of the members is enough to produce a relevant phenotype (Carapuço et al., 2005).

Another characteristic of vertebrate *Hox* genes is their temporal and spatial collinearity, that is, the order of these genes in the chromosome is the same as the order of their expression along their main body axis during development. The first ones to be expressed are the ones closer to the 3' region and are key pieces in the development of more anterior regions (Duboule, 1998; Duboule and Dollé, 1989).

### **I.2.3 The *Hox* code**

The *Hox* code is a theory that states that it is the different combinations of *Hox* gene expression at a particular axial level that specifies segmental identity in the anterior/posterior axis (Kessel and Gruss, 1991). This theory is supported by certain characteristics, such as the previously mentioned collinear expression and the fact that *Hox* genes have what is called a posterior prevalence or phenotypic suppression: basically, a more posterior *Hox* is functionally dominant over its anterior companions (Duboule and Morata, 1994). This means that there are spatial barriers and that the anterior borders of *Hox* expression become progressively more posterior for the 5' *Hox* genes.

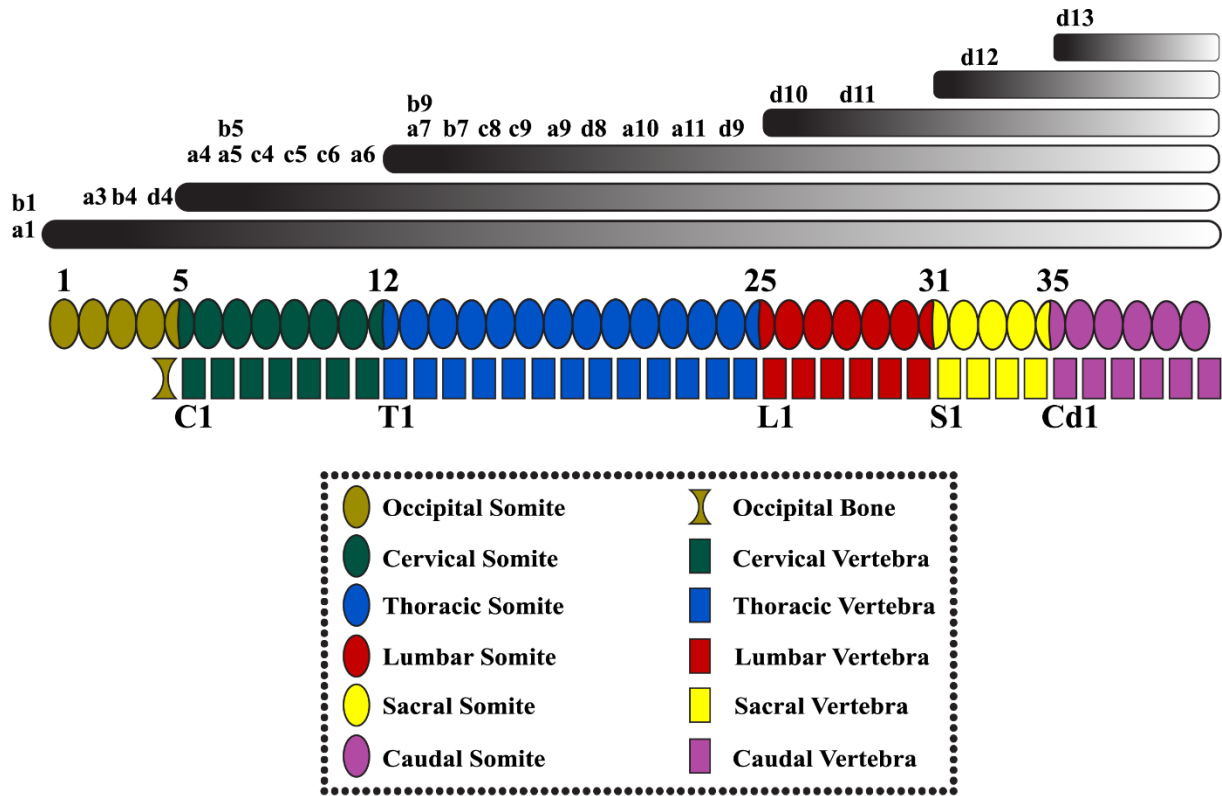
A good example of the *Hox* code is provided by the *Ultrabithorax* (*Ubx*) mutation. Usually expressed in the third thoracic segment (T3), *Ubx* inhibits wing formation (Lewis, 1978). Through a loss-of-function mutation, it was shown that an extra pair of wings developed in the thorax. *Antp* is normally expressed in T2 and responsible for wing development. In the absence of *Ubx* expression, *Antp* expands its functional domain into the T3, where *Ubx* is normally expressed. This allows *Antp* to extend its wing development effect one more segment, hence another pair of wings being formed. Since *abdA*, is still activated in the first abdominal segment (A1) and given the posterior prevalence of *Hox* genes, *Antp* ectopic expression stops and the normal structural development of the anterior/posterior axis continues after that point. For this reason, the *Ubx* mutant flies have two pairs of wings (Lewis, 1978).

### **I.2.4 Mouse *Hox* gene expression**

As mentioned above, the first *Hox* to be expressed are the ones located near the 3' region of the clusters. Before any somites are formed, *Hox* transcription starts in the primitive ectoderm of the embryo's posterior primitive streak and spreads in a rostral direction (Deschamps et al., 1999; Forlani et al., 2003). This is regulated by the same “players” as somitogenesis, WNTs, FGFs and retinoic acid (Deschamps et al., 1999). In more advanced stages, in the pre-somitic mesoderm, Notch and FGF signalling regulate *Hox* gene expression (Dubrulle et al., 2001; Zákány et al., 2001).

### I.2.5 *Hox* genes and the axial skeleton development

As mentioned above, there are five types of vertebral morphology in most vertebrates. This differentiation is largely dependent on *Hox* genes and their “code”. The different combinations of *Hox* genes expressed throughout the anterior/posterior axis are the key for the differentiation of somites to vertebrae of the correct type and in the correct place (Fig. 1.2) (reviewed in Mallo et al., 2010).



**Figure 1.2: Simplified representation of *Hox* gene expression domains during axial skeleton development.** The reducing expression gradient shows how *Hox* expression trends towards decrease in more posterior regions. The type of vertebra formed is labelled (C, Cervical; T, Thoracic; L, Lumbar; S, Sacral; Cd, Caudal). Adapted from Burke et al., 1995 and Favier and Dollé, 1997.

Paralogous groups *Hox1* through *Hox4* are the first groups to be transcribed. Located in the 3' extremity of the clusters, they are expressed in rostral regions. Two of these groups, *Hox3* and *Hox4*, have been associated with the correct patterning of the neck by having a role in cervical vertebrae morphogenesis (Condie and Capecchi, 1994; Horan et al., 1995). However, no *Hox* combination has been linked to a total change in the fate of the cervical region.

When we move more posteriorly, we enter the expression domain of *Hox5* to *Hox9*, which are essential for ribcage development (reviewed in Mallo et al., 2010). For example, it is well established that the *Hox6* group has rib-promoting activity, meaning that it can confer a thoracic identity to vertebral elements normally lacking ribs (Vinagre et al., 2010). This was proven by experiments where transgenic mice overexpressing *Hoxb6* in the presomitic mesoderm had ectopic ribs, extending throughout the lumbar region of the axial skeleton.

The *Hox10* group is responsible for generating the lumbar region. The first approach to check this group's role in the axial skeleton patterning was to inactivate all three *Hox10* genes in mice. The resulting *Hox10*-null mutant offspring had ribs along the thoracic and lumbar region and small ribs fused at their lateral margins of the sacral region (Wellik and Capecchi, 2003). In addition to this, overexpressing just one member of this group, *Hoxa10*, in the paraxial mesoderm gave origin to mice

with rib-less phenotypes (Carapuço et al., 2005). Together these experiments showed that the *Hox10* group has rib-repressing activity and, therefore it is essential for the thoracic to lumbar transition. This paralogous group will soon be explored in greater detail.

The *Hox11* group is essential for the correct formation of the sacral and caudal regions. In its absence, no sacrum is formed and when one of its members is overexpressed in the presomitic mesoderm, transgenic mice have “sacralisation” phenotypes (Carapuço et al., 2005; Wellik and Capecchi, 2003). These transgenics had the sacrum itself in a more anterior position and several fusions at the thoracic level, a characteristic of sacral vertebrae. As mentioned previously, *Hox10*-null mutants had in the sacral region small ribs fused at their lateral margins. Combined, this information indicates that *Hox11* partially suppresses *Hox10* and that *Hox10* also plays a role in sacral formation.

Finally, the *Hox13* group stops axial extension. The first experiments consisted in loss-of-function mutations, whereupon *Hoxb13* null mice showed slight overgrowth of the tails (Economides et al., 2003). Afterwards, transgenic mice overexpressing all the group’s members using the *Cdx* promoter had posterior truncation phenotypes. These mice had fewer caudal vertebrae and smaller tails (Young et al., 2009).

### **I.2.6 Hox protein function and DNA binding properties**

*Hox* proteins mainly bind DNA through their homeodomain (HD) (Pearson et al., 2005). However, the HD alone cannot explain specific functions, since different *Hox* proteins bind very similar target sequences (Noyes et al., 2008). In addition to this, it has already been shown that, for some *Hox* proteins, functional specificity is related to regions outside the HD (Galant and Carroll, 2002; Guerreiro et al., 2012; Ronshaugen et al., 2002). Even the same *Hox* protein can display different functions in different organisms resulting from distinct evolution of other regions of the protein. A good example is *Ubx*, which in *Drosophila* not only induces abdominal characteristics in the thoracic segments, but also inhibits limb formation. *Artemia*’s *Ubx* however, only induces abdominal characteristics. The reason for this is a transcriptional repression domain in *Drosophila*’s C- terminal region, which is absent in *Artemia* (Galant and Carroll, 2002; Ronshaugen et al., 2002).

In the mouse, similar results have been reported for *Hoxa11*. In this case, the authors substituted the *Hoxa11* HD with those of *Hoxa4*, *Hoxa10* and *Hoxa13*. The different constructs were all able to maintain some of *Hoxa11*’s functions, but produced developmental differences in other structures (Zhao and Potter, 2001; Zhao and Potter, 2002). This proves that, similarly to other groups, in vertebrates and, most particularly in mice, *Hox* functional specificity goes beyond the HD.

### **I.3 A molecular look at *Hoxa10***

Previously it was mentioned that the *Hox10* group has rib-repressing activity and plays a key role in the thoracic to lumbar transition (Wellik and Capecchi, 2003). It has also been mentioned that *Hox* genes have similar sequences and are sorted in paralogous groups according to their position and sequence similarities. Therefore, it is possible that conserved amino acid residues specific to each *Hox* group could explain, at least partially, the functional specificity of *Hox* proteins.

So far, the functional specificity of *Hox10* proteins has not been clearly explained. All three members of this paralogous group can block rib formation and no other group can mimic this activity (Carapuço et al., 2005; Wellik and Capecchi, 2003). Given their functional redundancy and because *Hoxa10* has been used before, it is a good representative of the *Hox10* paralogous group. This allows us the design of an experimental approach to get closer to understanding how it blocks rib formation (Carapuço et al., 2005; Guerreiro et al., 2012).

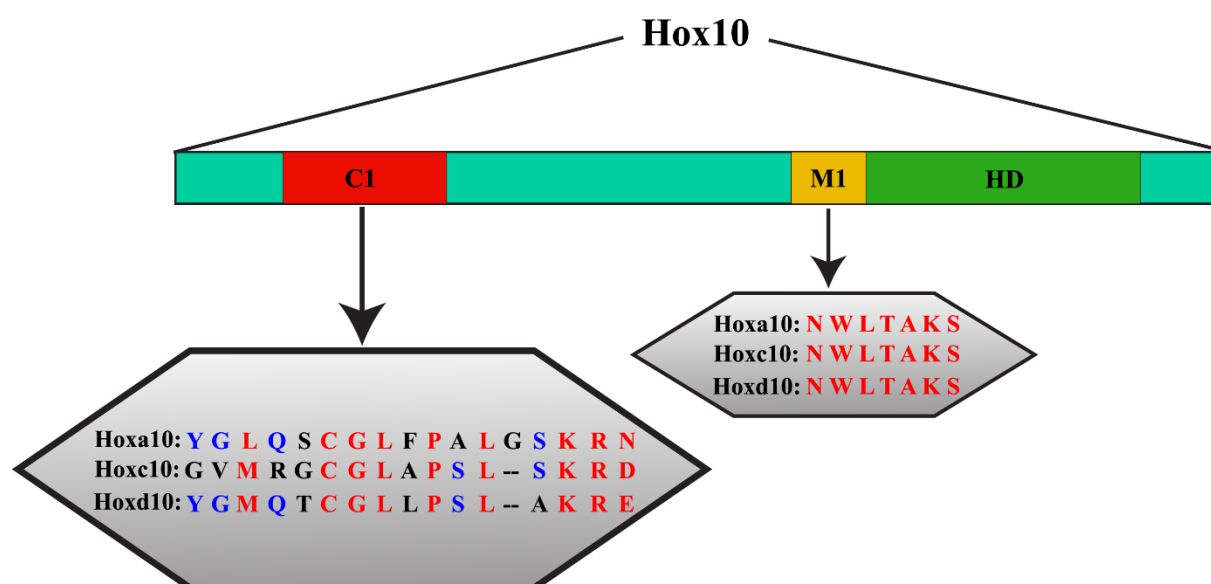
### I.3.1 Hoxa10's M1 motif

Previous work has been done trying to link unique Hoxa10 protein motifs and its rib-repressing ability. So far, only one motif has been identified and it was called M1 (Guerreiro et al., 2012). This is a conserved seven amino acid sequence located next to the N-terminal end of the homeodomain of Hox10. This motif has two phosphorylation sites, which are also vital for Hox10 activity. When these sites were mutated, the rib-repressing properties were lost.

However, despite its importance, the M1 motif is not sufficient for Hoxa10's rib-repressing activity. Transgenic mice overexpressing Hoxb9 containing the M1 motif of Hox10 proteins did not exhibit severe abnormal rib phenotypes, only small alterations in rib development (Guerreiro et al., 2012).

### I.3.2 Hoxa10's C1 motif

Unpublished work from Mallo's group indicates that Hoxa10's rib-repressing activity could be related to another conserved protein motif. Unlike M1, this motif is not completely conserved throughout the Hox10 group, however it is promising. It contains 17 amino acids and is referred to as C1 (Fig. 1.3).



**Figure 1.3: Schematic representation of Hox10 proteins and their conserved protein motifs.** The M1 motif and homeodomain sequences are identical among all Hox10 group members, but the M1 motif is not completely conserved.

The C1 motif is thought to act as a protein binding site. This hypothesis arises from related work that used this motif in a yeast two-hybrid system screen. Using this system, some proteins that contained WD40 domains interacted with C1. In addition to WD40-containing proteins, the yeast two-hybrid experiment identified Smad4 as another potential candidate for interaction with Hoxa10. This raises the question: is Hoxa10's function dependent, or at least, influenced by interactions of C1 with these proteins?

### I.4 Groucho-Related Gene

One of the proteins that interacted with the C1 motif in the previously mentioned experiment was Groucho-Related Gene 3 (Grg3), one of six members of the mouse Grg protein family (Leon and Lobe, 1997). Different organisms share this protein, but the nomenclature differs. In human, this is known as transducin-like enhancer of split (TLE) and in *Drosophila* as Groucho (Gro) (Jennings et al., 2008).

Grg3 is a transcriptional co-repressor that interacts with several DNA-binding repressors (Paroush et al., 1994). One example is the Hairy enhancer of split (Hes) and other Hairy-like proteins, which are known to interact with Grg proteins (or one of its homologues in other species) (Ju et al., 2004). These transcription factors are downstream effectors of the Notch signalling pathway (Kageyama et al., 2005). Since it is involved in the Notch pathway, Grg3 might play a role in several developmental processes, such as myogenesis, gut development, neurogenesis, somitogenesis and other cell determination processes (Bate et al., 1993; Fortini et al., 1993; Tepass et al., 1995). It has also been shown that Groucho plays an important role in *Drosophila* embryonic segmentation (Paroush et al., 1994).

Grg3 proteins contain a WD40 domain, a repetitive sequence of forty-four to sixty amino acids. Apart from its characteristic tryptophan-aspartic acid (W-D) dipeptide repeats, it also includes the glycine-histidine (GH) dipeptide (Neer et al., 1994; Smith et al., 1999). These domains function as a protein-protein or protein-DNA interaction platform, but mainly to serve as a rigid scaffold for protein-protein interaction (Stirnemann et al., 2010). Proteins containing WD40 domains are involved in many different cellular functions. These include, among others, apoptosis and transcription regulation (Li and Roberts, 2001).

It has been published that the WD40 domains are key players in the binding ability of Gro/TLE proteins to their targets (Jennings et al., 2006; Jime et al., 1997). In addition to this, *TLE3* expression has been reported in the sclerotome which in turn gives rise to the vertebrae and ribs (Dehni et al., 1995). This makes Grg3 a candidate to be involved in the development of the vertebral column, including the rib formation program.

*Grg3* expression also fits in the available time window for an interaction with *Hoxa10*. It has an expression pattern overlapping with the window of active Notch signalling during somitogenesis and it is found in the presomitic mesoderm (Leon and Lobe, 1997). Likewise, *Hoxa10* is expressed and active in the presomitic mesoderm (Carapuço et al., 2005). Also, *Grg* genes are known to be expressed during the somitogenesis phase of avian embryonic development (Van Hateren et al., 2005).

## **I.5 Smad4**

Smad4 was another of the possible *Hoxa10* binding partners identified in the yeast two-hybrid screen with the C1 domain. Smad is the name used for the vertebrate homologues of *Sma* and *Mad* (*Mothers against decapentaplegic*), first described in *Caenorhabditis elegans* and *Drosophila*, respectively (Savage et al., 1996; Sekelsky et al., 1995). These intracellular proteins are part of the transforming growth factor-beta (TGF- $\beta$ ) pathway, in which they are responsible for signal transduction by forming complexes that act as transcription factors (Derynck et al., 1998).

There are pathway-restricted Smads and common-mediator Smads. The first type is composed by Smad proteins that interact with specific receptors, either activin type I receptors or Bone morphogenetic protein (BMP) type I receptors (reviewed in Heldin et al., 1997). On the other hand, there is Smad4, the common-mediator Smad. Smad4 forms hetero-oligomers with pathway-restricted Smads and the resulting complex is translocated into the nucleus where it acts as a transcription factor (reviewed in Heldin et al., 1997). There are also Smads with an inhibitory function (Heldin et al., 1997).

Smad4 can bind directly to DNA, however, it has no transcriptional activity by itself (Derynck et al., 1998). Previous studies in mice have revealed some importance of Smad4 in gastrulation, but also in the anterior/posterior axis patterning (Sirard et al., 1998). In addition, *Smad4* is expressed in several embryo regions including the presomitic mesoderm (Gray et al., 2004). These characteristics make Smad4 another interesting candidate for *Hoxa10* functional interactor.

## **I.6 Objectives**

Currently, it is known that *Hox10* genes have a rib-repressing activity (Carapuço et al., 2005; Vinagre et al., 2010; Wellik and Capecchi, 2003) and that this ability is partially explained by the M1 and C1 protein motifs (Guerreiro et al., 2012 and unpublished work). This work, intends to expand this knowledge by identifying novel proteins that have the molecular potential to be interacting with Hoxa10, most particularly proteins that have WD40 domains, which act as a possible interaction site and proteins that are involved in the anterior/posterior axis patterning. These interactions were studied using a human cell line, in order to have a system similar to *in vivo* surroundings. In addition, this work also aims to understand the potential role of the C1 motif in Hoxa10 function, as well as its role in the potential interaction with the previously mentioned candidate proteins.

---

## II. Materials & Methods

---

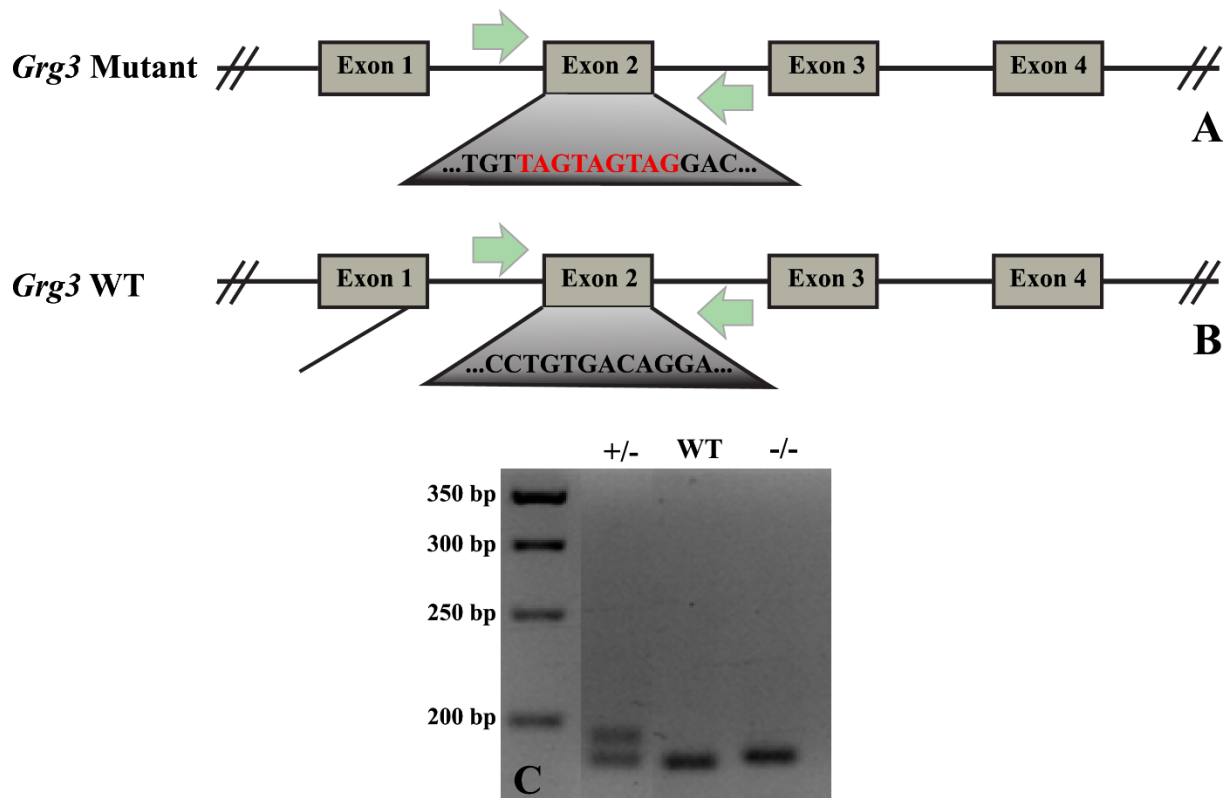
### II.1 Animal model

For this work, specific pathogen-free mice of the FVB or C57Bl/6 strains were used. These animals were kept at room temperature (RT) on 12 hours dark/light cycles. They were used for transgenic, mutant and wild type (wt) analyses.

### II.2 Making *Grg3* knockout mutants

To generate a new mouse line without an active *Grg3* the CRISPR/Cas9 system was used (Yang et al., 2013). First, oocytes were injected with a mixture of: 1) a gRNA targeting a sequence within *Grg3*'s second exon (Appendix III); 2) the Cas9 protein; and 3) a single stranded DNA oligonucleotide containing three stop codons designed to stop the *Grg3* reading frame, flanked at each side by sixty nucleotides of sequence homology to the *Grg3* sequence (Appendix III). These homologous sequences are those around the area targeted by the gRNA. The injected oocytes were then transferred into the oviduct of pseudo pregnant females. The microinjection process will soon be explained in slightly greater detail.

#### II.2.1 Genotyping *Grg3* mutants



**Figure 2.1: Genotyping strategy of *Grg3* mutants.** A) *Grg3* mutant allele contained 3 tandem stop codons (in red) in frame with the *Grg3* protein, placed in exon 2. Location of the primers used for genotyping is indicated with the green arrows. B) *Grg3* wild type sequence in the targeted region of exon 2. C) Electrophoretic analysis of PCR fragments corresponding to a wild type (WT), a heterozygous (+/-) and a homozygous (-/-) mutant for *Grg3*.

The animals born from these injections were genotyped to identify those carrying the mutant allele. For this, a small piece of the mouse tails was cut and placed in tail lysis buffer (Appendix II) containing 200 µg/ml of Proteinase K (pK) overnight (ON) at 50°C. The pK was then inactivated at 100°C for ten minutes and 1 µl of the solution used as template for a polymerase chain reaction (PCR). The primers used for genotyping are shown in Appendix III (3322 and 3325). Mutants had extra nine nucleotides, which corresponded to three added stop codons. The genotyping strategy is represented in Figure 2.1 A and B and the agarose gel analysis (Appendix I) is shown in Figure 2.1 C. The heterozygous samples appeared to have two bands, but in fact they were four. The upper band was a result of hybrid chains (wild type and mutant copies annealing together) and the lower band are actually two bands: the wild type DNA (190 bp) and the mutant DNA (199 bp).

Since this first offspring was mostly chimeric in terms of *Grg3* allelic composition, those that carried the mutation were crossed with wild-type mice to obtain an offspring with a clean *Grg3* allelic composition (F1 - Filial 1). The genotyping procedure was then repeated.

Heterozygous males and females were crossed and this originated the F2. The resulting homozygous embryos were used to assess the impact of absent *Grg3* expression in axial skeleton patterning and, especially, in the rib formation program.

### **II.3 Making transgenic constructs**

To make transgenic constructs standard molecular biology techniques were used. These constructs were used for different purposes, which will be explored later.

#### **II.3.1 Constructs used in this work**

The different constructs used for microinjection, were already inserted in the pBluescript® II KS+ vector. All these constructs contained a *Dll1* promoter (a presomitic mesoderm promoter) (Beckers et al., 2000) a FLAG tag octapeptide and the *Hoxa10*'s 3'UTR and polyA tail. The different *Hoxa10* and *Hoxb9* molecules were inserted in frame with the FLAG tag and upstream of the 3'UTR.

For some transfection experiments, the cDNAs were inserted in the multiple cloning site of the pCMV Sport6.1 vector, which contained either a c-MYC tag or a FLAG tag peptide.

#### **II.3.2 cDNA synthesis**

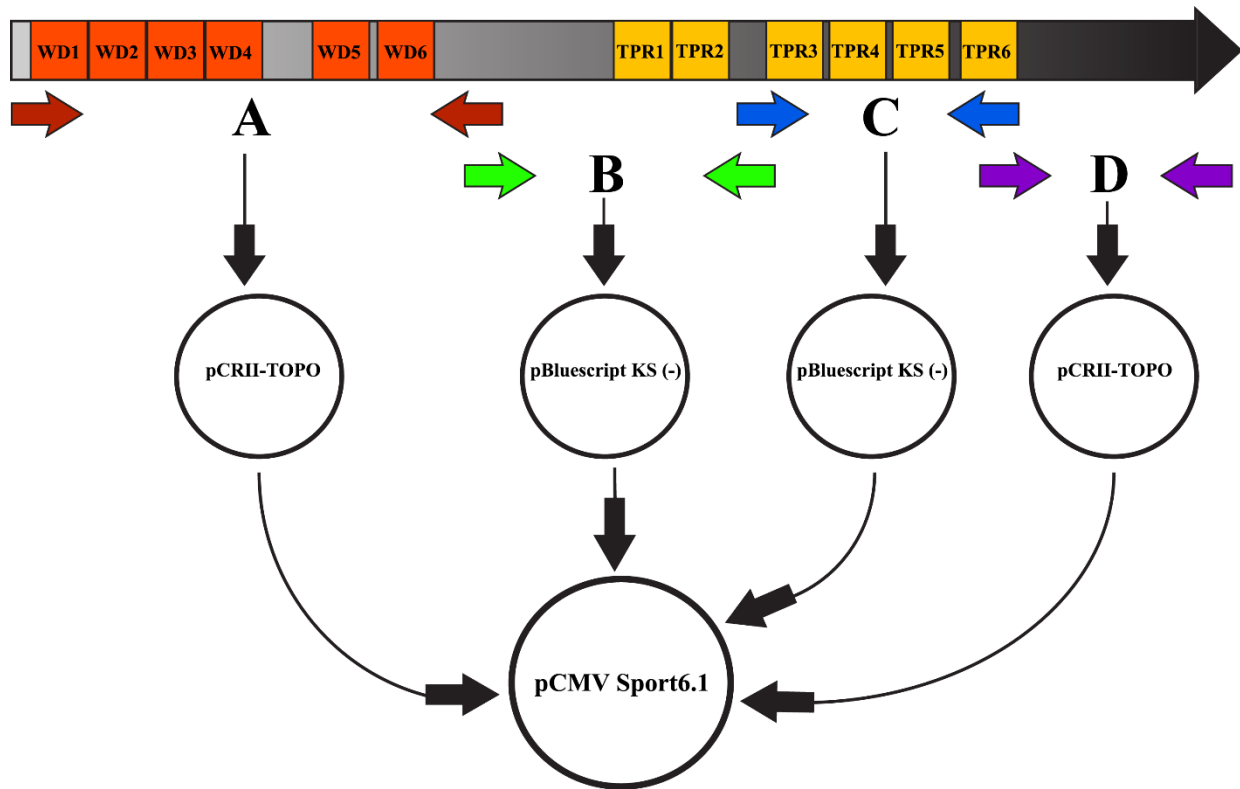
To obtain cDNAs that were not available or to generate probes for *in situ* hybridization, they were synthesized using the Reverse Transcriptase synthesis followed by PCR. Embryonic day (E) 8.5 and 10.5 mouse embryo total RNA was used as template. A 17 µl mixture containing RNA (1 µg), dNTP's (1 µl, stock at 10 mM), random hexamers (1 µl, stock at 250 ng/µl) and H<sub>2</sub>O was incubated at 65°C for 5 minutes and then, placed on ice for one minute. At this point, 2 µl of Reverse Transcriptase 10X buffer and 1 µl of Reverse Transcriptase enzyme were added to the mixture and incubated at 25°C for 10 minutes, 50°C for 50 minutes and at 85°C for 5 minutes. The newly synthesized cDNA was stored at -20°C and ready to be amplified by PCR to obtain the relevant cDNAs.

For *Grg3* only its WD40 domain was amplified by PCR using primers 3348 and 3349 (see Appendix III). The PCR product was purified by phenol/chloroform extraction, followed by ethanol precipitation. The resulting DNA was digested with Bam H I and Not I.

Production of *IFT144* cDNA was performed sequentially given its large size (around 4.2 kb). As seen in Figure 2.2, the coding sequence was divided in four portions (A, B, C and D) and specific primers for each portion were used (Appendix III). Each PCR product was then purified by phenol/chloroform extraction and precipitated using ethanol. With this step completed, each fragment



was inserted into a vector and transformed. The different pieces were then assembled into the Sport6.1 vector.



**Figure 2.2: IFT144 cloning strategy.** Four different overlapping portions were amplified from embryonic RNA by RT-PCR and individually cloned into an intermediary plasmid. They were later assembled together in the Sport6.1 plasmid (containing a c-MYC-tag). The different combinations of primers used for each portion (A, B, C and D) are represented in the figure with different colours and the primer sequences are described in Appendix III.

### II.3.3 Plasmid digestion, isolation and purification for microinjection

To obtain the transgenic construct for injection, the vector sequences were removed by digesting the corresponding plasmid (20 µg) with SpeI, XbaI and PvuI (1 µl each) in a final 50 µl volume at 37°C ON.

The digested DNA was then run in a 1% agarose gel in 1X TAE with 5 µl of Ethidium bromide (Appendix II). The band corresponding to the transgenic construct was cut out from the gel. The DNA was purified from the agarose using the NZYGelpure kit according to the manufacturer's protocol and eluted in 50 µl of the kit's elution buffer. The concentration of the purified DNA was measured using NanoDrop 1000 Spectrophotometer and stored at -20°C.

Intermediate molecular procedures such as ligation, transformation, colony screening by PCR, digestion analysis, DNA purification and Sanger sequencing are described in detail in Appendix I.

### II.4 Microinjection

Transgenic embryos were generated by pronuclear injections. After hormone priming, superovulated females were mated with males and fertilized oocytes retrieved. The previously obtained DNA fragments were then microinjected into the pronuclei at a concentration of 2 ng/µl. The injected oocytes were later transferred into the oviduct of pseudo pregnant females, a procedure performed by the IGC's Transgenic Unit.

## **II.5 Embryo analysis**

First, wild type embryos were collected from pregnant females by extracting the uterus and placing it in 1x PBS. Afterwards, the embryos were separated from the decidua. Once isolated, they were placed in 4% Paraformaldehyde (PFA in PBT) ON at 4°C. The day after, embryos were dehydrated through five minute washes with increasing concentrations of methanol/PBT and stored in 100% methanol at -20°C.

When working with transgenic or mutant embryos it was also necessary to genotype each one.

### **II.5.1 Genotyping**

For E8.5 to E12.5 embryos, yolk sacs were collected and placed in 50 µl of yolk sac lysis buffer (Appendix II) containing pK (100 µg/ml) and incubated at 50°C overnight. pK was inactivated at 95°C for 10 minutes and 1 µl of this solution was then used in a PCR reaction with the relevant primers (Appendix III).

For E13.5 to E18.5 embryos their intestines were used for DNA extraction. They were incubated in 500 µl of Laird's buffer containing pK (200 µg/ml), at 50°C ON with shaking. To precipitate the DNA, 500 µl of isopropanol were added and the DNA "fished" with a pipette tip into 250 µl of TE, and dissolved at 37°C ON with shaking.

The PCR conditions are shown in Appendix II.

### **II.6 Skeletal staining**

This procedure was performed with embryos from E13.5 to E18.5. Embryos were skinned, eviscerated, and fixed in 100% ethanol. For staining, embryos were incubated in alcian blue staining solution (Appendix II) for 12 to 20 hours (RT) and then fixed in ethanol 100% ON. Soft tissue was then digested using 2% Potassium Hydroxide for 6 hours at RT, followed by a 3-hour incubation in an alizarin solution red (Appendix II) and further incubation ON in 1% Potassium Hydroxide. Tissue digestion was stopped in 25% glycerol and the cleared embryos stored at RT.

## **II.7 *In situ* hybridization**

### **II.7.1 RNA probe synthesis**

Antisense RNA probes for the genes to be analysed were synthesized by *in vitro* transcription. The probes used in this work are listed on Appendix III.

10 µg of vector DNA containing the probe was first linearized using the appropriate restriction enzyme, followed by Phenol/Chloroform purification (Appendix I) and ethanol precipitation. The resulting pellet was resuspended in 10 µl of H<sub>2</sub>O to be used for the transcription reaction. Transcription was performed using T3 or T7 RNA polymerase, in order to obtain the anti-sense strand. The transcription reactions were performed on RNA-free tubes and contained 0,5 to 1 µg of template, 0,6 µl of RNase inhibitor, 2 µl of 10x digoxigenin-labelled nucleotides (Roche), 2 µl of 10x transcription buffer 1 µl of T3 or T7 RNA polymerase and H<sub>2</sub>O to complete 20 µl. The reaction was incubated at 37°C for 1 hour. Afterwards, 1 µl was run on an agarose gel to confirm that the transcription reaction occurred. The probe was then precipitated by adding 80 µl of TE pH 8.0, 10 µl of 3M NaOAc, pH 5.2 and 250 µl of 100% ethanol, leaving it in dry ice for 30 minutes. The probe was recovered by centrifugation (14000 rpm at 4°C), the supernatant removed, the pellet air dried and resuspended in 80 µl of TE pH 8.0.

### II.7.2 *In situ* procedure

All the following steps were performed at room temperature, except when mentioned otherwise.

**Table 2.1:** Proteinase K incubation times for each embryonic stage.

Embryonic Stage	pK Incubation Time
<b>E8.5</b>	4 minutes
<b>E9.5</b>	7.5 minutes
<b>E10.5</b>	9 minutes
<b>E11.5</b>	10 minutes

1<sup>st</sup> Day: Embryos were rehydrated with decreasing concentrations of methanol in PBT and then incubated in PBT. They were bleached with 6% hydrogen peroxide (diluted in PBT) for 1 hour, with rotation. Afterwards, they were washed 3 times in PBT (5 minutes each) and treated with pK (10 mg/ml). The incubation time depended on the embryo stage (see Table 2.1).

To stop pK activity, the embryos were washed with a freshly made glycine solution (2 mg/ml in PBT) for 5 minutes, followed by 3 washes in PBT (5 minutes each) and re-fixed in 4% PFA (in PBT)/0.2% glutaraldehyde for 20 minutes. After further washes in PBT, the embryos were incubated with 0,5 ml of prewarmed hybridization solution 1 (see Appendix II). After the embryos sank, 2 ml of the prewarmed same solution were added and incubated 1 hour at 65°C. Embryos were then incubated with 1 ml of prewarmed hybridization solution 1 containing 6 µl of the specific mRNA probe at 65°C ON.

2<sup>nd</sup> Day: The hybridization solution 1 with the probe was removed and the embryos were washed twice with prewarmed hybridization solution 2 (Appendix II), at 65°C (30 minutes each). The washes continued in prewarmed TBST/ hybridization solution 2 (Appendix II) at 65°C for 30 minutes and with three additional washes with TBST, 5 minutes each and with MABT, three times (5 minutes each). Afterwards, embryos were incubated in 10% blocking solution (Appendix II) for two and half hours. Finally, the blocking solution was substituted with 1% blocking solution containing Anti-Digoxigenin-AP antibody (1:2000) and the embryos were left incubating at 4°C, ON (with rocking).

3<sup>rd</sup> Day: Embryos were washed three times with MABT (Appendix II), 5 minutes each and then washed five times with MABT, 1 hour each. They were left in MABT ON.

4<sup>th</sup> Day: Embryos were washed three times with NTMT (Appendix II), 5 minutes each. Afterwards, the alkaline phosphatase reaction was developed by incubation with developing solution (Appendix II) in the dark. The completeness of the development was checked in the microscope. To stop the reaction, the embryos were washed twice in PBT (5 minutes each) and fixed with 4% PFA at 4°C ON (with rocking). The next day, they were stored in PBT.

### II.8 Cell culture

293T cells, which are human embryonic kidney cells transformed with the large T antigen (stock in liquid nitrogen at -170°C), were thawed at 37°C and diluted in 5 ml of feeder medium (Appendix II). This mixture was centrifuged for 5 minutes, 1000 rpm (RT), the supernatant removed and the cell pellet resuspended in 5 ml of feeder medium. These 5 ml were plated in a 60 mm dish and incubated at 37°C in a 5% CO<sub>2</sub>-containing atmosphere. Cells were passed when they reached around 95% confluence. For this, cells were washed twice with 2 ml of PBS (Dulbecco's phosphate buffered saline). Then, 0,8 ml of prewarmed Trypsin-EDTA (0.25%; Appendix II) were added and incubated at

37°C for 5 minutes. Trypsin was inactivated by diluting the cell suspension in 5 ml of feeder medium, the suspension was centrifuged at 1000 rpm for 5 minutes and the cell pellet resuspended in an amount of feeder medium that depended on the size of the dishes. 2 ml were used for 35 mm dishes and 5 ml for 60 mm dishes.

### **II.8.1 Transfection**

Cell medium was substituted for transfection medium (Appendix II) and the cells incubated for 4 hours at 37°C in a 5% CO<sub>2</sub> containing atmosphere. The transfections were performed using Lipofectamin™ 2000 following the manufacturer's instructions. Depending on the dish area, either 4 µg (35 mm dish) or 8 µg (60 mm dish) of plasmid DNA were used. In this case, different constructs inserted in the pCMV-Sport6.1 were used. Transfected cells were incubated at 37°C in a 5% CO<sub>2</sub>-containing atmosphere for about 24 hours.

### **II.8.2 Cell lysis and storage of cell extracts**

Cells were washed with PBS, scrapped in 1 ml of PBS, using a scrapper, the suspension was transferred to a 1.5 ml eppendorf tube and kept on ice. The suspension was centrifuged for 1 minute, 6500 rpm, at 4°C. The pellet was resuspended in 40 µl of nondenaturing lysis buffer (Appendix II) and left on ice for 1 hour. This is a moderate-strength lysis buffer effective for whole-cell extracts (including nuclear proteins), which does not disrupt protein complexes. Afterwards, the mixture was centrifuged for 20 minutes, at 4°C (14000 rpm) and the resulting supernatant (the cell extracts) transferred to a new tube and stored at -80°C.

### **II.8.3 Protein extract analysis**

The detection of tagged proteins in the cell extract was done by Western Blot.

First, 10 µl of each sample was boiled in 6x SDS loading buffer (Appendix II) for 5 minutes. The samples were separated in a SDS-PAGE gel (either 8% or 12% depending on the predicted size of the proteins; Appendix II) containing a 5% SDS polyacrylamide stacking gel (Appendix II). The gel was run at 110 V for around two hours. Proteins were then Blotted into a PVDF membrane (Polyvinylidene difluoride), previously equilibrated in methanol and water, in transfer buffer (Appendix II) at 200 mA, for 1 hour. Membranes were then soaked in 5% blocking solution (Appendix II) for another two hours (RT). Afterwards, blocking solution (Appendix II) was removed and the membranes were incubated with primary antibody (1:1000 in blocking solution) ON, at 4°C.

The next day, the membrane was washed three times in PBT and incubated in the secondary fluorophore-conjugated antibody, either anti-mouse or anti-rabbit (1:10000 in 5% blocking solution) for 1 hour (RT) and then washed three times in PBT.

Signal detection was completed using Odyssey Scanner, which generates images from excitation and emission of fluorescent molecules in the near infrared range.

## **II.9 Co-Immunoprecipitation (Co-IP)**

After confirming that the extracts had the tagged proteins, a co-immunoprecipitation approach was used in order to detect if Grg3's WD40 domains, IFT144 and Smad4 proteins interacted *in vivo* with Hoxa10.

1st Day: 10 µl of Dynabeads® Protein G (Appendix II) were equilibrated in 100 µl of lysis buffer for 10 minutes, at room temperature (with rocking). After removing the supernatant (using a magnet), these beads were used for lysate preclearing. For this, 30 µl of lysate were incubated with the

Dynabeads® for around six hours, at 4°C (with rocking). This step is important in order to reduce non-specific binding to beads. The supernatant (the pre-cleared lysate) was transferred to a new tube and kept on ice.

Since the different transfected constructs contained a FLAG or c-MYC tag, both antibodies could be used as baits. Given that these Dynabeads® did not contain either one, there was a need to create a complex between the beads and the antibody. 25 µl of bead suspension were washed 10 minutes in PBT, RT with rocking. After the supernatant was removed, to couple Dynabeads® with the antibody (either anti-FLAG or anti-c-MYC) the beads were resuspended in 200 µl of PBT containing 5 µg of antibody. As a control, the same procedure was done, but adding non-immune IgG (mock Ip). The mix was incubated for 6 hours, RT with rocking). The supernatant was then removed and the Dynabeads® were washed in PBT.

After the removal of PBT from the Dynabeads®/antibody complex, 30 µl of cleared lysate was added to the Dynabeads® coupled with either anti-FLAG or anti-c-MYC and IgG as a control and incubated at 4°C, RT with rocking.

To improve efficiency, when using anti-FLAG antibodies for the IP, an alternative protocol was also used with Anti-FLAG M2 magnetic beads instead of Dynabeads®. These are 4% agarose magnetic beads covalently bound to an Anti-FLAG mouse antibody. In this case, there was no need for preclearing or for a mock Ip. These beads were first equilibrated in 100 µl of lysis buffer for 20 minutes, RT with rocking. 30 µl of cell extracts were then added to the beads. The rest of the protocol was the same as when using Dynabeads®.

2nd Day: After the overnight incubation, the supernatant was recovered and stored for later analysis. This supernatant contained everything that was not directly or indirectly captured by the antibody.

The Dynabeads®/antibody/antigen complex was washed two times, using 200 µl of Washing Buffer (Appendix II). The beads were then resuspended in 100 µl of the same buffer and transferred to a clean tube. For the elution, the supernatant was removed and the Dynabeads® resuspended in 6x SDS loading buffer (Appendix II), boiled 5 minutes and centrifuged 15 min at 14000 rpm (4°C). The supernatant was retrieved and this is used for further Western Blot analyses.

## **II.10 Immunostaining**

This was done to determine the subcellular localization of the Grg3-WD40 domain and the IFT144 protein. For the Grg3-WD40 domain, 293T cells were transfected with the corresponding expression plasmids. In the case of the IFT144 protein, it was the endogenous protein what was detected.

Immunostaining was performed in cells grown on glass coverslips.

1<sup>st</sup> day: After transfection and incubation for 24 hours, cells were then incubated in PBS-T 0,3% (Appendix II) during 5 minutes at RT. They were then washed twice in PBS and blocked with PBS containing 10% FBS (Fetal Bovine Serum; Appendix II) for 1 hour at RT. Afterwards, cells were washed three times with PBS containing 0,5% FBS and incubated with primary antibody, diluted in a similar buffer at 4°C ON. Antibody concentrations are shown in Appendix II.

2<sup>nd</sup> day: Cells were washed 5 times in PBS-FBS 0,5% (Appendix II), followed by incubation with the secondary antibody (Appendix II) diluted in a similar buffer, for 1 hour at RT. Cells were washed again in PBS-FBS 0,5%, 3 times and incubated in a DAPI solution in PBS-FBS 0,5% for 5

minutes. The coverslips were mounted on a slide, using a small drop of VECTASHIELD® Mounting Medium (Appendix II). This was then sealed using nail polish and stored at 4°C.

Images were acquired using a Leica DMRA2 upright microscope, coupled to a CoolSNAP HQ CCD camera. Image analysis was performed using Fiji (Schindelin et al., 2012).

---

## III. Results

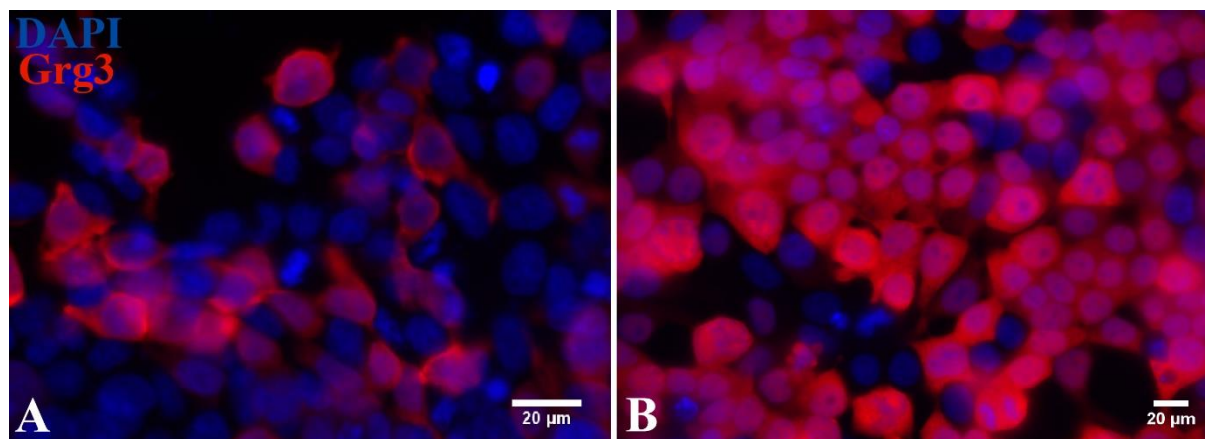
---

### III.1 Hoxa10 and its functional candidates

#### III.1.1 Grg3

##### III.1.1.1 Grg3 Co-IP

To determine if Grg3 could interact with Hoxa10 in the context of mammalian cells, 293T cells were transfected with constructs expressing a FLAG-tagged Hoxa10 and the c-MYC tagged WD40 domain of Grg3 (cMYC-WD40), which, in theory, is responsible for its interaction with other proteins. The initial cMYC-WD40 construct was not localized in the nucleus (Fig. 3.1A), most likely because it did not contain a nuclear localization signal (NLS). As this could thus hinder interaction with Hoxa10, it was necessary to add a NLS to this construct.



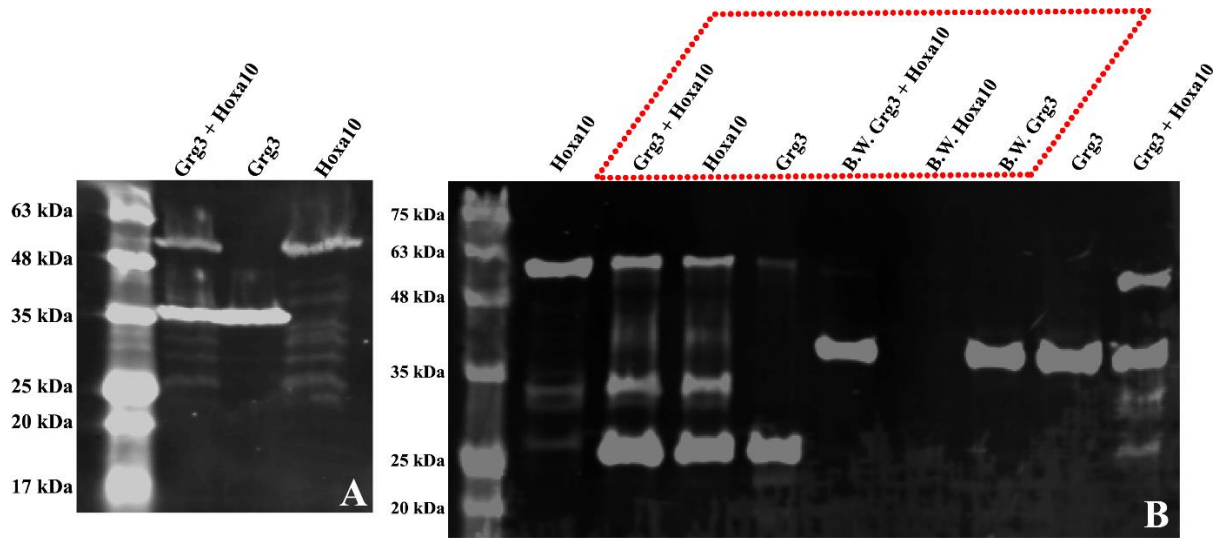
**Figure 3.1: Immunostaining of Grg3-WD40 transfected 293T cells.** A) The Grg3-WD40 construct lacking an NLS did not appear to be localized in the nucleus, being visible mostly in the cytoplasm. B) The Grg3-WD40 construct containing the NLS was present in the nucleus. Grg3-WD40 was detected with anti-c-MYC, the nuclei stained with DAPI and the images captured using the 40x 0.75NA objective on a Leica DMRA2 microscope.

With an additional cloning step, the Simian virus 40 (SV40) NLS was introduced to the construct on the C-terminus, just after the cMYC tag, and immunostaining analysis revealed the presence of this protein in the nucleoplasm (Fig. 3.1B).

After transfection and analysing cell extracts by Western Blot to confirm the presence and sizes of the tagged proteins (Fig. 3.2A), Co-IP experiments were performed. In theory, when precipitating the target protein, interacting proteins will also be present in the immunoprecipitated and identified in a Western Blot using antibodies against the tags of the transfected proteins.

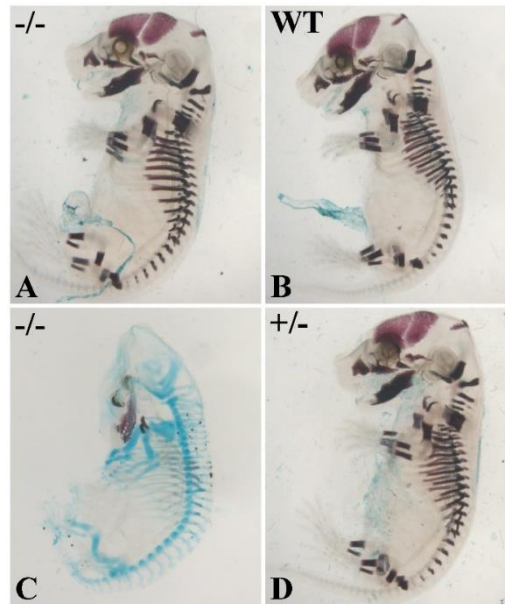
Analysis extracts from cells transfected with the Grg3 and Hoxa10 constructs revealed the presence of these molecules, although their molecular weight was slightly different to the expected 37,2 kDa and 42,3 kDa, for Grg3-WD40 and Hoxa10, respectively (Fig. 3.2A). As seen in Fig. 3.2B, Grg3 was not detected in the immunoprecipitated complex obtained with the anti-FLAG antibody, indicating that it does not interact with Hoxa10. The presence of large amounts of Grg3 in the supernatant recovered from the immunoprecipitated complex (referred as B.W. in Fig. 3.2B) further supported this conclusion. In addition, the absence of Hoxa10 in the non-bound samples indicate that it was effectively immunoprecipitated by the anti-FLAG antibody. Co-IP experiments using Grg3-WD40 as bait produced

equivalent results (not shown). Therefore, when taking all the results into account, we conclude that Hoxa10 and Grg3 do not interact, at least not in 293T cells.



**Figure 3.2: Grg3-WD40 Co-IP analysis.** **A)** The presence of Grg3-WD40 (c-MYC) and Hoxa10 (FLAG) in cell extracts of transfected 293T cells was assessed by a Western Blot using antibodies against the respective tags. The cells had been transfected with Grg3-WD40 and Hoxa10, alone or in combination. **B)** The Co-IP results using Hoxa10 as bait. The region marked with the red dotted box shows the immunocomplexes obtained with extracts of cells transfected with Hoxa10 and Grg3-WD40 alone or in combination, and the corresponding supernatants of the immunocomplexes (B.W.) stained with antibodies against the two tags. The protein inputs are shown in the lanes outside the red box. Grg3-WD40 was not detected in the immunoprecipitated complex that contained Hoxa10, but was detected in the supernatant. Hoxa10 was detected in the immunoprecipitates from cells transfected with this protein but not in the supernatant.

### III.1.1.2 Grg3 Knockout mice



**Figure 3.3: Alcian blue and alizarin red staining of Grg3 KO embryos.** **A)** Homozygous E15 *Grg3* knock-out embryo. **B)** Wild type E15 embryo. **C)** Homozygous E13 *Grg3* knock-out embryo. **D)** Heterozygous E15 *Grg3* knock-out embryo. None of the mutant embryos had any obvious morphological alteration, except a smaller size. However, homozygous E13 *Grg3* knock-out embryos seemed to lack ossified ribs due to delayed development (C).

*Grg3*'s knockout (KO) experiments had been reported by another group investigating its role in placenta development (Gasperowicz et al., 2013). Those experiments showed that most of *Grg3*-deficient embryos died before E15.5, but no embryonic analysis was performed. We wanted to



determine whether these embryos had skeletal defects and created a mouse line mutant for *Grg3* using the CRISPR/Cas9 system (Yang et al., 2013). In our experiments, we analysed E13 and E15 embryos for bone and cartilage staining (Fig. 3.3).

At first glance, E13 homozygous KO *Grg3* embryos seemed to lack ossified ribs as we could only detect cartilage in the corresponding area (Fig. 3.3C). However, E15 embryos clearly showed rib development, although with a possible developmental delay (Fig. 3.3A). Indeed, *Grg3* mutant embryos were globally less developed but without obvious morphological differences when compared to wild type littermates. The only visible variation appeared to be their smaller size. This was probably caused by their placental defects and therefore, these embryos might be in early stages of developmental arrest.

To confirm these results, E10.5 embryos were used for *in situ* hybridization. The chosen probe was *Myf5* (*Myogenic Factor 5*), a muscle marker induced after somite formation, because it is under *Hoxa10* regulation during rib formation (Vinagre et al., 2010). Only one mutant was recovered at E10.5, with the others being either wild type or heterozygous. As seen in Figure 3.4, the expression pattern of *Myf5* remained unaffected, as well as its intensity, which is in keeping with the lack of skeletal defects.



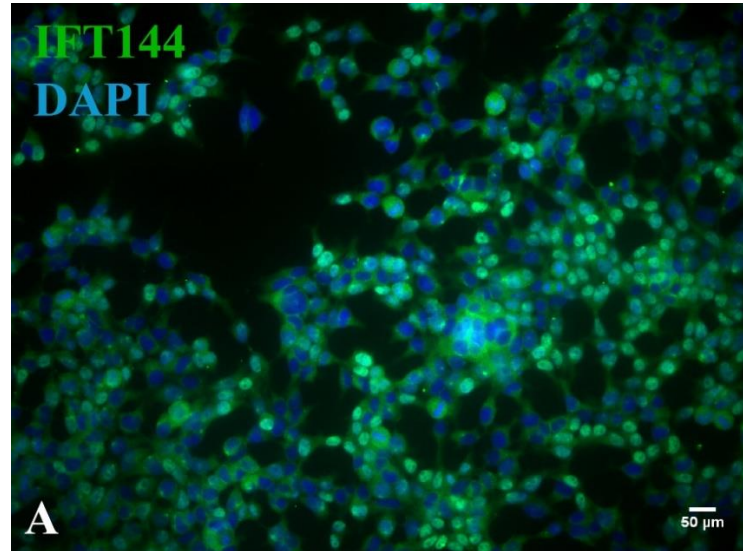
**Figure 3.4: *In situ* hybridization of *Grg3* KO E10.5 embryos.** E10.5 mouse embryos were stained for *Myf5* expression, a muscle marker under *Hoxa10* regulation during rib formation. There is no difference in *Myf5* expression between homozygous null mutant (A), heterozygous (B) and wild type (C) embryos, indicating that *Grg3* does not influence somitogenesis.

### III.1.2 IFT144

The above results indicate that *Grg3* might not be a physiological cofactor of *Hoxa10*. As several other WD40 domain-containing proteins were identified in the yeast two-hybrid screen from the Mallo laboratory, we thought that the physiological *Hoxa10* cofactor could be another WD40-containing protein, including those that did not come out in the screen. We therefore searched for WD40-containing proteins that had a reported mutant phenotype compatible with *Hoxa10* functional activity. IFT144 was one such molecule as its inactivation led to defects in rib development and early somite patterning (Ashe et al., 2012).

Intraflagellar transport proteins (IFT) mediate the trafficking system responsible for transporting proteins required for cilia assembly and function (Pedersen et al., 2008). Proteins that mainly regulate retrograde transport (from the cilium tip to the cell body) are part of the IFT-A complex and proteins that mainly mediate anterograde transport form the IFT-B complex (Ocbina et al., 2011). Several skeletal dysplasia diseases have been associated to this family of proteins. To our interest, the IFT-A gene *IFT144* (also mentioned in some literature as *WDR19*) has been correlated to patients diagnosed with the Jeune and Sensenbrenner syndromes. These are ciliopathies that cause short ribs and limbs, polydactyly and craniofacial defects (Bredrup et al., 2011). IFT144 not only possesses WD40 domains, but also tetratricopeptide repeats (TPR). This is another structural motif that forms scaffolds to mediate protein–protein interactions (Blatch and Lässle, 1999). Despite being a protein involved in ciliary transport, images from the datasheet information of a commercial antibody indicate that it is probably present in the nucleoplasm, an essential characteristic for a *Hoxa10* interactor. We confirmed the nuclear

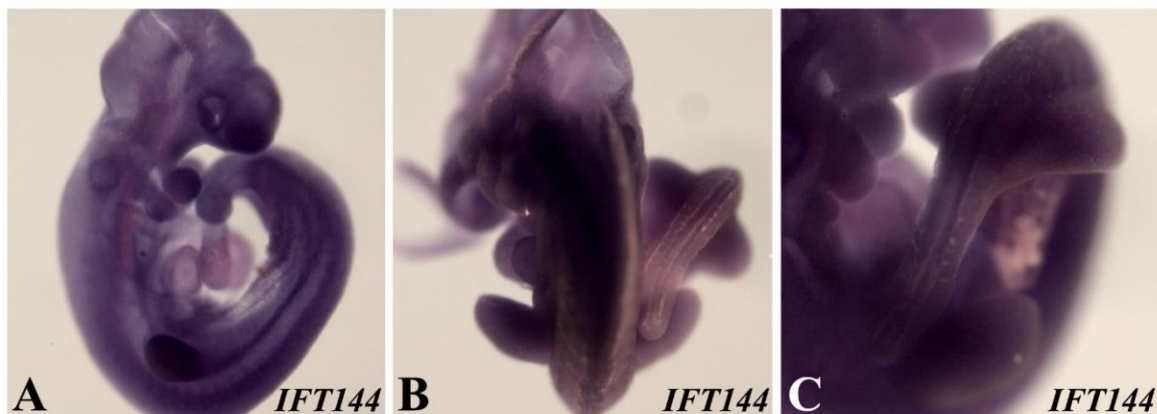
localization of IFT144 using a cell immunostaining protocol. In this experiment, no transfections were performed, since human cells contain endogenous IFT144 recognized by a commercial antibody. Results showed that IFT144, much like *Hoxa10*, is localized in the nucleus (Fig. 3.5). This observation, together with the IFT144 mutant phenotype further makes this protein a potential candidate to interact physiologically with *Hoxa10*.



**Figure 3.5: Immunostaining of 293T cells for endogenous IFT144.** IFT144 was localized in the nucleus, as well as in the cytoplasm. Protein detection was performed using a commercial antibody for IFT144, the nuclei stained with DAPI and the images captured using the 10x 0.25NA objective.

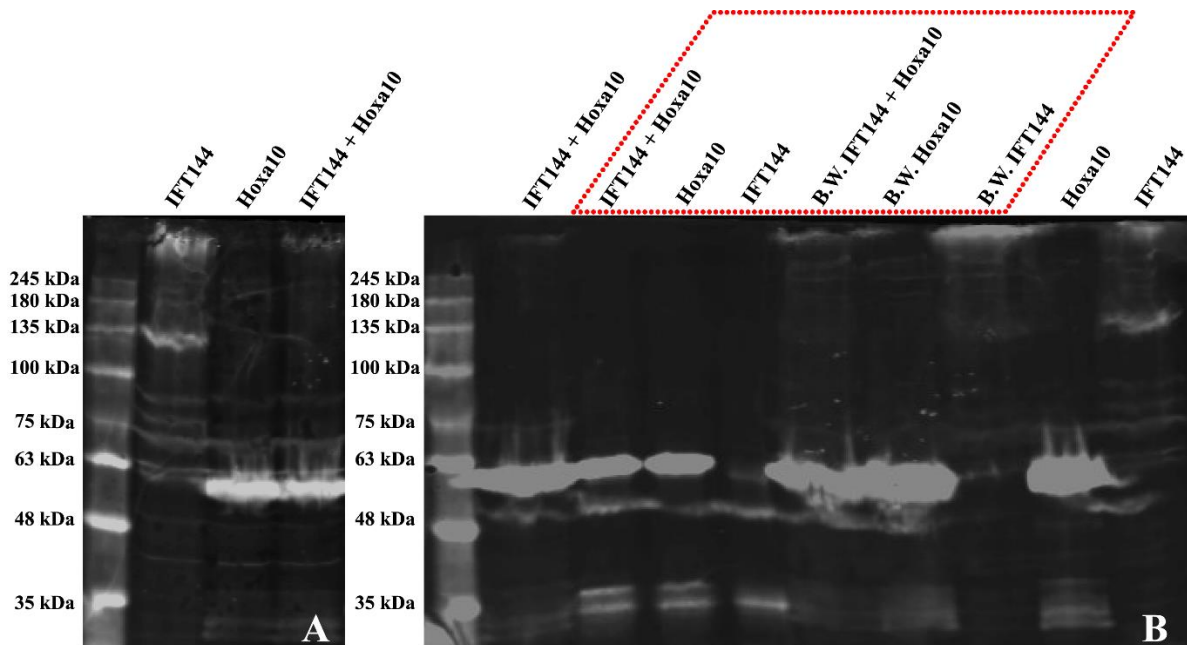
To understand whether IFT144 could interact with *Hoxa10*, transfection experiments like the ones with *Grg3* were performed. In addition, *in situ* hybridizations in wild type embryos were also performed to map this gene's expression.

*IFT144*'s expression pattern is shown in Figure 3.6. No clear distinct pattern was observed for this probe. It is possible that some of the signal is background, however, *IFT144* appears to be ubiquitously expressed in the wild type embryos, including in the presomitic mesoderm. Ideally, another probe targeting a different region of *IFT144* would have been used to confirm these observations, but time constraints did not allow it.



**Figure 3.6: *In situ* hybridization showing *IFT144* expression.** The RNA probe targeted the least conserved region of among IFT genes. Despite plenty of background, it seems that IFT144 is ubiquitously expressed in wild type embryos at E9.5 (A) and E10.5 (B, C).

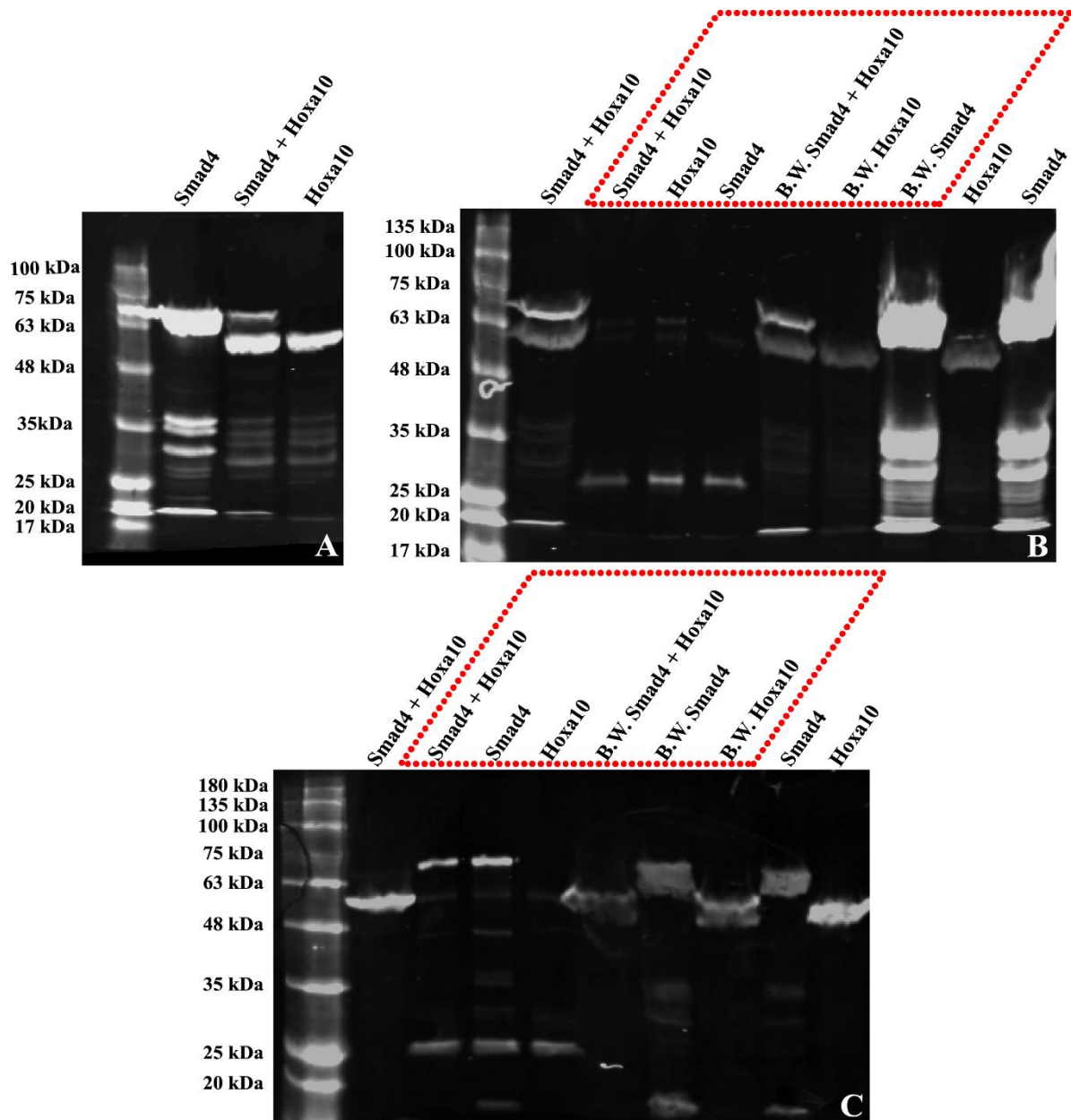
Since *IFT144* is found in the nucleus, a Co-IP was performed using extracts from cells transfected with a vector expressing *Hoxa10* (containing a FLAG tag). In theory, IFT144 should be detected in Western Blots since there is endogenous IFT144, as seen above (Fig. 3.5). However, despite using the same commercial antibody for IFT144, we could not detect this protein in cell extracts by Western Blot. For this reason, we produced the full coding sequence of mouse *IFT144* and tagged it with c-MYC, using as template E8.5 and E10.5 mouse embryo cDNA. Given that this was a very large sequence (around 4.2 kilobases), the sequence was divided in four pieces, taking advantage of the several restriction enzymes found in the coding region. The individual portions were then cloned together in the Sport6.1 expression vector and this was transfected into 293T cells together with the *Hoxa10* expressing vector. However, it is worth mentioning that the observed size for the tagged IFT144 did not match the predicted 152,8 kDa in the Western Blot of the transfected cell extracts that would later serve as input for the Co-IP. When IFT144 was transfected alone, we could detect two bands (around 135 kDa and over 245 kDa) that were not observed in the non-transfected cells and in the double IFT144-*Hoxa10* transfected cells only the 245 kDa was apparent (Fig. 3.7A). Surprisingly, these bands were detected with the commercial antibody against IFT144, whilst anti-c-MYC showed no signal. The origin of this paradoxical result could be that the cDNA in the transfected construct contained two potential start codons, one before c-MYC and one after (the natural IFT144 start codon) and, therefore, it is possible that the c-MYC tag was not expressed. Nevertheless, given the clear difference in protein molecular weight it is possible that the two bands do not represent IFT144, although it is clear that they should be somehow related to it, given that they are only present in extracts from cells transfected with IFT144. Even though these conditions were not the best, we used these extracts for the Co-IP experiments. Once again, both *Hoxa10* and IFT144 were used as baits to guarantee that the observed results were consistent.



**Figure 3.7: IFT144 Co-IP analysis.** **A)** The presence of IFT144 (c-MYC) and *Hoxa10* (FLAG) was assessed by a Western Blot in cell extracts from 293T cells transfected with IFT144 and *Hoxa10* (single and co-transfection). *Hoxa10* was detected using anti-FLAG and IFT144 using a commercial antibody. **B)** Analysis of the immunocomplexes and supernatants from the Co-IP experiments using *Hoxa10* as bait. Immunocomplexes together with the corresponding supernatants (B.W. labelled lanes) are shown inside the red dotted box and the inputs used are shown in the lanes outside of the box. IFT144 was not detected in the immunoprecipitated complex, but it is detected in the supernatant (B.W. IFT144 + *Hoxa10*). The membranes were stained with antibodies against both FLAG and IFT144.

When analysing the immunoprecipitates done with the anti-FLAG antibody on the double transfected extracts, despite intense background, none of the IFT144-related bands seemed to co-elute with Hoxa10 (Fig. 3.7B), as no signal was observed when the IP was tested with anti IFT144 antibody. The 245 kDa band was, however, detected in the supernatant. A similar result was observed with the control extracts not containing Hoxa10. Together, these results suggest that no interaction occurred between Hoxa10 and IFT144, although better controlled experiments for the actual presence of IFT44 should be performed.

### III.1.3 Smad4



**Figure 3.8: Smad4 Co-IP analysis.** A) Western Blot of cell extracts of 293T cells transfected with constructs expressing Smad4 (FLAG) and Hoxa10 (c-MYC) (single and co-transfection). They were detected using antibodies against both tags. B) Analysis of the immunocomplexes and supernatants from the Co-IP experiments using Hoxa10 as bait. Immunocomplexes together with the corresponding supernatants (B.W. labelled lanes) are shown inside the red dotted box and the inputs used are shown in the lanes outside the box. Smad4 was not present in the immunoprecipitated complex (Smad4 + Hoxa10 sample) and was detected in the supernatant (B.W. Smad4 + Hoxa10 sample). This indicates that Smad4 and Hoxa10 do not interact. Proteins were detected using antibodies against the two tags. C) A similar experience as in B) but using Smad4 as bait.



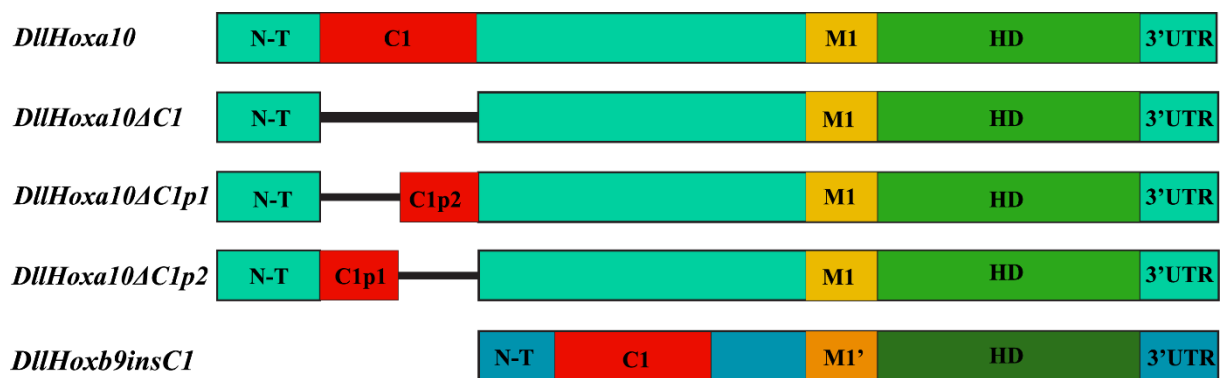
Smad4 was explored as a final potential functional interactor for Hoxa10. For this protein, only transfections and Co-IPs were performed. In this particular case, the human *Smad4* coding sequence was already inserted in a vector containing a FLAG-tag and therefore in these experiments we used a construct expressing a Hoxa10 protein containing a c-MYC-tag. Transfected cells for the tagged Smad4 contained a protein of the predicted molecular weight of this protein (61.4 kDa) that reacted to antibodies against the tag in Western Blots (Fig.3.8).

Once again, both proteins were used as baits in the Co-IP. The supernatants recovered before the first wash were also used for protein detection, as well as the eluted product of the Co-IPs performed with extracts from cells transfected with Hoxa10 or Smad4 alone. As seen in Figure 3.8 (B and C), both Western Blot revealed that Smad4 was not co-eluted with Hoxa10, which means that these two proteins do not seem to interact in cultured mammalian cells. It should, however, be noted that in the Co-IP where Hoxa10 was used as bait, the Hoxa10 signal is rather weak (Fig. 3.8B), and therefore, the amount of this protein could be too low to pull down detectable levels of Smad4. The results using Smad4 as a bait are stronger, but Smad4 was not detected in the Co-IP input lane (first lane in Fig. 3.8C). However, this was probably just a detection error since Smad4 is present in the Co-IP output of both co-transfected and Smad4 only extracts (second and third lane in Fig. 3.8C). In addition, Smad4 was also detected in the Western Blot, using the collected extracts to confirm both tagged proteins (Fig. 3.8A). Much like IFT144, despite not having the best resolution and quality, Hoxa10 and Smad4 do not appear to be interacting, at least at a protein-protein level.

### III.2 The C1 motif and the axial skeleton

As mentioned above, Hoxa10 possible interactions and its rib-repressing activity seems to be influenced by the C1 motif. To confirm its impact in axial patterning, several *Hoxa10* deletion mutants and a *Hoxb9*-chimeric construct were over expressed in the presomitic mesoderm of mouse embryos.

The constructs used were *DllHoxa10ΔC1*, *DllHoxa10ΔC1p1*, *DllHoxa10ΔC1p2* and *DllHoxb9insC1* and are represented in Figure 3.9. *DllHoxa10ΔC1* consists in the Hoxa10 sequence after deleting the entirety of the C1 motif. The second and third constructs have the C1 motif partially deleted: *DllHoxa10ΔC1p1* lacks the first eight C1 amino acids and *DllHoxa10ΔC1p2* lacks the last eight. Finally, *DllHoxb9insC1* has the C1 motif inserted into Hoxb9, which has no intrinsic rib-repressing properties.



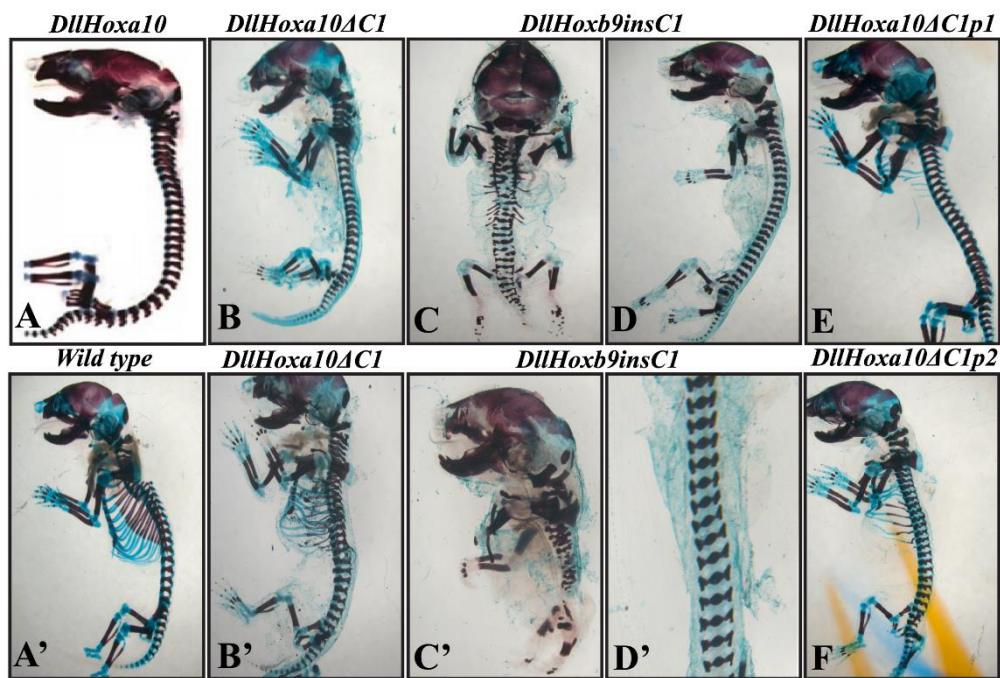
**Figure 3.9: Schematic representation of each transgenic construct used in the microinjections.** N-T corresponds to the N-terminal region, which contained a FLAG-tag in all constructs. *DllHoxa10ΔC1* completely lacked the C1 motif and *DllHoxa10ΔC1p1* and *DllHoxa10ΔC1p2* lacked parts of this motif. The *DllHoxb9insC1* contains the native Hoxb9 sequence, with the C1 motif added into the N-terminal part of the protein.

A total of five E18.5 *DllHoxa10ΔC1* transgenics were analysed. Three of them had the wt phenotype, one had a mild phenotype (lacking some ribs) and one was completely rib-less (Fig. 3.10 B and B', respectively). When *Hoxa10* is overexpressed in the presomitic mesoderm, rib-less phenotypes

are common (Fig. 3.10A) and, importantly, milder phenotypes are not observed (Carapuço et al., 2005). Therefore, even considering the small number of embryos analysed, these results suggest that the C1 motif is necessary for *Hoxa10* rib-repressing functions, since that activity is, at least, partially lost.

E18.5 transgenic embryos derived from the *DllHoxa10ΔC1p1* and *DllHoxa10ΔC1p2* constructs were also helpful. One out of four of the *DllHoxa10ΔC1p1* transgenics was completely rib-less and the other three had a wild type phenotype (Fig. 3.10E). *DllHoxa10ΔC1p2* did not result in any embryos with a complete rib-less phenotype: 5 out of 7 had normal axial skeleton phenotypes, one was almost rib-less (Fig. 3.10F) and one had a mild rib-less phenotype (lacking one pair of ribs, not shown). Once again, these results reveal a role for the C1 motif in axial skeleton development, since *Hoxa10*'s rib-repressing ability is reduced.

To understand if the C1 motif was sufficient to confer rib-repressing activity the *DllHoxb9insC1* construct was injected, and twelve E18.5 transgenic embryos were recovered and stained. 8 of those embryos had rib-less phenotypes (Fig. 3.10 D), as well as clear segmentation defects in the axial skeleton, including truncations and fused vertebrae (Fig. 3.10 C, C' and D'). Two had a wild type phenotype and the remaining two had mostly segmentation problems with most ribs absent (Fig. 3.10 C and C'). These results clearly show that C1 is sufficient to confer rib-repression to Hox proteins lacking this activity. In addition, the results also show that C1 could interact with the vertebrae segmentation processes.



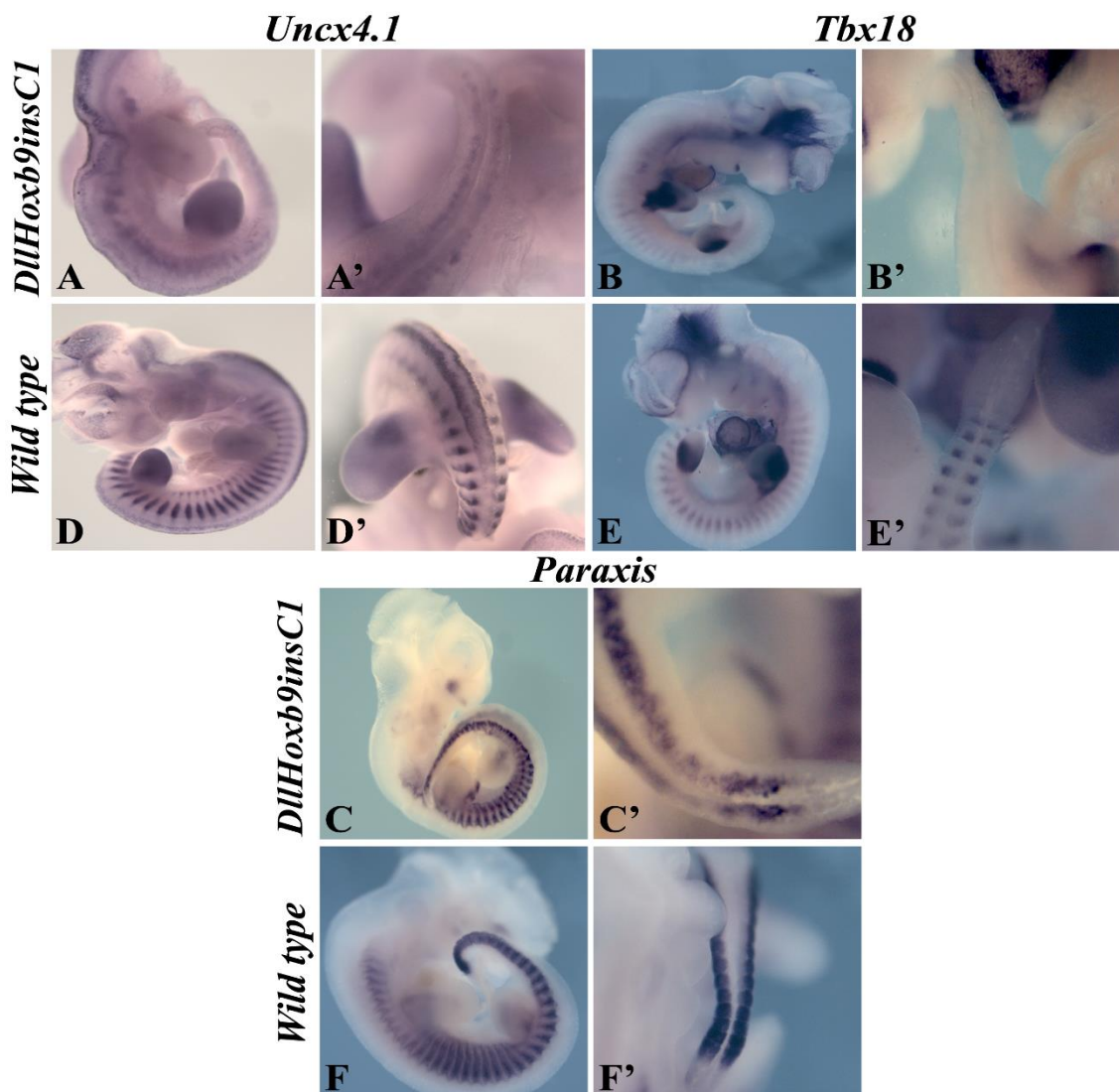
**Figure 3.10: Representative phenotypes of E18.5 embryos overexpressing different transgenic constructs.** A) Typical rib-less phenotype when the *DllHoxa10ΔC1* construct is overexpressed, adapted from Carapuço et al., 2005. A') Wild type embryo. B, B') E18.5 *DllHoxa10ΔC1* embryos had mostly wild type phenotypes. Some exceptions were either rib-less (B), or contained mild phenotypes (B') lacking some ribs. C, C', D, D') *Dllb9insC1* embryos had rib-less phenotypes (D and D'). They also showed severe segmentation problems, such as truncation and more subtle ones, like fused vertebrae (C and C'). E) *DllHoxa10ΔC1p1* transgenics had normal phenotypes, but one had a rib-less phenotype, with no segmentation issues (E). F) A *DllHoxa10ΔC1p2* embryo exhibiting an intermediate rib phenotype without any segmentation defects. The majority of these embryos had wild type phenotypes.

By analysing these E18.5 transgenic mice it became obvious that the C1 motif has a role in *Hoxa10* protein function, and it might affect vertebrae segmentation as well. Poorly formed vertebrae are most likely a consequence of defects in the somite patterning stage. To see if this was indeed the case, constructs associated with segmentation problems were microinjected and embryos were collected

at E10.5 and E11.5. *DllHoxb9insC1* transgenic embryos were analysed by *in situ* hybridization, using *Uncx4.1*, *Tbx18* and *Paraxis* as molecular markers (Fig. 3.11).

Not all transgenic embryos showed a morphological phenotype, which is in keeping with the lack of 100% penetrance of the skeletal phenotype, and therefore, we focused on those that had obvious morphological alterations in the somites when that was the case.

*DllHoxb9insC1* E11.5 transgenic embryos, stained for either *Tbx18* or *Uncx4.1* showed an obviously disrupted pattern of expression. These are markers of the anterior and posterior somite halves, respectively. *Tbx18* expression was absent in the one transgenic embryo with a visible somitic phenotype, especially in the more posterior somites (Fig. 3.11 B and B'). Similarly, *Uncx4.1* expression was absent in the more posterior somites and not properly localized in the more anterior somites (Fig. 3.11 A and A'). The degree of disruption was not the same in all embryos, but it seems that more posterior somites were generally more affected. Together, both probes show that the C1 motif can somehow interfere with the somitogenesis process.



**Figure 3.11: *In situ* hybridizations of E11.5 *DllHoxb9insC1* transgenic embryos with somitogenesis markers.** A, A') *Uncx4.1* expression was absent from the more posterior somites of *DllHoxb9insC1* transgenic embryos (A') and out of place in more anterior somites (A), when compared to the normal expression pattern (D and D'). B, B') *Tbx18* expression was also absent in these transgenics, especially in the more posterior part of the embryo (B'). Wild type embryos had *Tbx18* expression restricted to the anterior half of the somite (E and E'). C, C') *Paraxis* expression was also affected, with no clear individual elements being formed, particularly in the more posterior regions.

Since somite patterning was clearly affected in these transgenics, an additional *in situ* hybridization was performed using *Paraxis*, another somitogenesis marker. E11.5 transgenic embryos showed a disruption of the expression pattern of this marker (Fig. 3.11 C and C'). Since *Paraxis* is essential for somite epithelization, this is an indication that no individual segments were properly formed from the presomitic mesoderm. Once again, the C1 motif's effect seems to disturb the development of the more posterior somites.



---

## IV. Discussion

---

This work had two main objectives: 1) to find proteins that functionally interact with Hoxa10, therefore influencing its rib-repressing activity; 2) understanding what role the C1 motif plays in axial patterning.

Our results clearly show that Grg3 does not interact with Hoxa10. Despite using the least stringent conditions possible, including a non-denaturing lysis buffer, Grg3 was never present in the complexes pulled down with an antibody detecting Hoxa10 in co-transfected cell extracts (Fig. 3.2). It is possible that the Grg3 construct used did not guarantee all of the interaction properties of the protein, given that only the WD40 domain was used. However, this was unlikely the reason, since these domains act as a rigid scaffold for protein-protein interaction and they have been reported as an important piece in Gro/TLE protein binding to other proteins (Jennings et al., 2006; Jime et al., 1997; Stirnimann et al., 2010). Given that the Grg3 construct also includes a SV40 NLS, both proteins were localized in the nucleus, a variable confirmed by the immunostaining assay (Fig. 3.1B). Therefore, the absence of interaction between Grg3 and Hoxa10 was not caused by a spatial impediment.

In addition to the apparent lack of interaction with Hox10, analysis of *Grg3* null mutants indicated that this gene is not involved in axial patterning, further supporting that it might not be a *bona fide* functional Hoxa10 interactor. Our data showed that these mice had a normal skeleton, including their ribcage. The only observed morphological differences in the mutants were the smaller embryo sizes, as well as being globally less developed (Fig. 3.3). The reason for this might be that *Grg3* mutants have severe placental defects (Gasperowicz et al., 2013). Since the placenta is responsible for the transport of nutrients and oxygen into the foetus, their smaller size is not surprising or the fact that most of the extracted homozygous mutants were in early stages of developmental arrest. The *in situ* hybridization data provided additional evidence for the lack of mutant phenotypes involving the ribcage (Fig. 3.4). E10.5 embryos had no differences in *Myf5* expression, which is under the positive regulation by Hoxa10 during rib development (Vinagre et al, 2010). Together, all the evidence is congruent. Grg3 does not interact with Hoxa10, and therefore it is also reasonable to suggest that it also does not bind other members of the Hox10 paralog group. In fact, Grg3 does not seem to have any functions related to rib development, or for that matter, axial skeleton development.

One of the other candidates for an interaction with the Hox10 proteins was IFT144 because it contains a WD40 domain and its mutant phenotype included rib defects (Ashe et al., 2012). As Fig. 3.6 showed, IFT144 expression is ubiquitous. This result was not surprising, since IFT144 is a key piece in cilia formation and function and, therefore, it could be expressed in many different tissues. Interestingly, endogenous IFT144 protein was detected in the cytoplasm and in the nucleoplasm of 293T cells (Fig. 3.5). Both results contributed for a generally better characterization of this gene and protein, since no published and peer-review work had shown any evidences for its expression or protein location.

Since both results were encouraging, IFT144 was overexpressed in 293T cells together with Hoxa10 to check if they were functional cofactors. In this case, the interaction was favoured by the fact that the vector contained the entire coding sequence, instead of just its protein-protein interaction domains (WD40 and TPR). However, it is not clear whether IFT144 was actually expressed in the transfected cells since the bands obtained with the antibody against IFT144 were too different from the predicted size (Fig. 3.7), although their presence was only detected in transfected cell extracts. As mentioned above, *IFT144* null mutants showed many axial patterning defects, including somite

patterning abnormalities, disordered arrangement of the vertebrae and rib defects (Ashe et al., 2012). So, even if IFT144 would turn out not to interact with the Hox10 proteins, it is still necessary for the correct development of the axial skeleton. This may be a by-product of IFT144 influence in cilia function, which could be upstream of axial development or could result from its involvement in a more specific function in somite differentiation. IFT proteins have been shown to be critical for activation and transduction of the Hedgehog (HH) signalling (Singh et al., 2015) and it has been reported that *Sonic hedgehog* (*Shh*) is required for the correct timing of somite formation (Resende et al., 2010). Also, IFT mutations have revealed similar phenotypes to those caused by deficient HH signalling, including shortened ribs which are associated with deficient Indian hedgehog (*Ihh*) signalling (Ashe et al., 2012; St-Jacques et al., 1999). Therefore, it is possible that these phenotypes are a consequence of poor HH signalling (including *Shh*), caused by the effect of IFT proteins in the position and activity of HH receptors in cilia. This idea of IFT proteins acting as a regulator of HH signalling in general, and *Shh* signalling in particular, is supported by the literature on this subject (reviewed in Goetz et al., 2009).

The final candidate for a possible interaction with *Hoxa10* was *Smad4*. This was the only case where a human sequence was used for the constructs. However, much like with the other tested candidates, we were unable to detect interactions between the two proteins (Fig. 3.8). Even though *Smad4* plays an important role in gastrulation and in anterior/posterior patterning, *Smad4*-deficient embryos actually have a posterior region with properly formed somites and defects mostly in the anterior region (Sirard et al., 1998). One can assume that, not only *Smad4* does not interact with *Hoxa10*, but given the reported phenotypes of *Smad4*-deficient embryos, it is also not involved in posterior embryo patterning.

The second main objective of this work comprised in unveiling the role played by the C1 motif in axial patterning and development. For this, microinjections of different transgenic constructs were performed. Unpublished results from the Mallo laboratory had already hinted at the possibility of this motif having an impact in *Hoxa10* function, as well as in segmentation. The E18.5 *DllHoxa10ΔC1* transgenics are consistent with this idea (Fig. 3.10 B and B'). This construct lacks the C1 motif and, in most embryos, the rib-repressing activity of *Hoxa10* was lost or reduced. If this motif was not important, overexpressing this construct would produce the same or similar results to those observed when *Hoxa10* is overexpressed, meaning more rib-less phenotypes (Carapuço et al., 2005). This result by itself indicates that the C1 motif is necessary for the rib-repressing activity of *Hoxa10*, although it might have some redundancy from other still non-identified part(s) of the molecule. In addition, we used two constructs that contained partial deletions of the C1 motif (Fig. 3.10 E and F). In both, the majority of the transgenic E18.5 embryos had wild type phenotypes, with few incomplete or complete rib-less embryos. These results were in total agreement with the *DllHoxa10ΔC1* transgenics, showing once again, that the C1 motif is necessary for a properly functioning *Hoxa10*. The *DllHoxb9insC1* transgenics provided more evidences for the role on C1 in rib formation, as it was able to provide rib-repressing properties to *Hoxb9*, a protein that by itself does not affect the rib formation program (Fig. 3.10 C and D). Most of the E18.5 embryos had rib-less phenotypes. Therefore, the C1 motif is not only necessary for the functionality of the Hox10 group, but it is sufficient to induce rib-repressing functions in the context of a Hox protein. However, C1 activity seems to go beyond the rib formation program. It also seems to have a role during somitogenesis, hence the abnormal vertebrae morphologies (Fig. 3.10 C, C' and D'). This becomes evident when analysing E11 *DllHoxb9insC1* transgenics, which originated segmentation problems (Fig. 3.11). All three segmentation markers revealed clear defects in early somitogenesis. The pattern of *Paraxis* expression in the paraxial mesoderm was disorganized indicating that somite segments were not properly formed (Fig. 3.11 C and C'). Additionally, *Uncx4.1* and *Tbx18* expression was mostly gone (Fig. 3.11 A, A', B and B', respectively). Both are markers for anterior/posterior somite patterning. In all three markers, the posterior region of the embryos was the

most affected by the atypical expression patterns. Therefore, the C1 motif somehow has the ability to interact with the segmentation network. When this motif is placed in its normal context (Hoxa10), it does not affect somitogenesis. However, when introduced in a non-physiological protein context, like the one provided by Hoxb9, it does. As mentioned before, unpublished results of the Mallo group showed that *DllHoxa10ΔC1p2* embryos have segmentation defects as well. It is possible that altering the C1 motif also affects its molecular context, revealing other properties, usually concealed. This was an interesting and surprising result, as a role in segmentation is not among the known functions of Hox genes. However, a rather similar result has been reported using Hoxb6 (Casaca et al., 2016). In particular, alterations in the linker region of Hoxb6 led to evident segmentation problems. Together, these results indicate that the rib-repressing activity of Hoxa10 could have some impact in the segmentation network. This effect occurs before somite patterning and probably just affects the segmentation network that is restricted to the sclerotome (the precursor of vertebrae and ribs). Also, given that the more anterior somites are not as affected as the more posterior, this effect could have a time window. Analysing transgenic embryos overexpressing a construct containing C1 at earlier development stages could be helpful.

In conclusion, Grg3, IFT144 and Smad4 seem to not functionally interact with Hoxa10 and, possibly also not with the other Hox10 group proteins. As for the C1 motif, it is not only necessary, but it is actually sufficient for Hoxa10 rib-repressing functions. This conclusion could be extended to the rest of the Hox10 group as it is one of the very few conserved regions across Hox10 proteins. The C1 motif might also play an unknown role in early somitogenesis, by somehow interacting with the segmentation network associated with the sclerotome.



## References

- Ashe, A., Butterfield, N. C., Town, L., Courtney, A. D., Cooper, A. N., Ferguson, C., Barry, R., Olsson, F., Liem, K. F., Parton, R. G., et al. (2012). Mutations in mouse *Ift144* model the craniofacial, limb and rib defects in skeletal ciliopathies. *Hum. Mol. Genet.* **21**, 1808–1823.
- Bate, M., Rushton, E. and Frasch, M. (1993). A dual requirement for neurogenic genes in *Drosophila* myogenesis. *Development* **119**, 149–161.
- Beckers, J., Caron, A., Hrabé de Angelis, M., Hans, S., Campos-Ortega, J. A. and Gossler, A. (2000). Distinct regulatory elements direct *Delta1* expression in the nervous system and paraxial mesoderm of transgenic mice. *Mech. Dev.* **95**, 23–34.
- Blatch, G. L. and Lässle, M. (1999). The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *BioEssays* **21**, 932–939.
- Bredrup, C., Saunier, S., Oud, M. M., Fiskerstrand, T., Hoischen, A., Brackman, D., Leh, S. M., Midtbø, M., Filhol, E., Bole-Feysot, C., et al. (2011). Ciliopathies with Skeletal Anomalies and Renal Insufficiency due to Mutations in the IFT-A Gene *WDR19*. *Am. J. Hum. Genet.* **89**, 634–643.
- Brent, A. E. and Tabin, C. J. (2002). Developmental regulation of somite derivatives: muscle, cartilage and tendon. *Curr. Opin. Genet. Dev.* **12**, 548–557.
- Brent, A. E., Schweitzer, R., Tabin, C. J., Ahrens, P. B., Solursh, M., Reiter, R. S., Aoyama, H., Asamoto, K., Beresford, B., Borycki, A.-G., et al. (2003). A Somitic Compartment of Tendon Progenitors. *Cell* **113**, 235–248.
- Burke, A. C., Nelson, C. E., Morgan, B. A. and Tabin, C. (1995). Hox genes and the evolution of vertebrate axial morphology. *Development* **121**, 333–46.
- Carapuço, M., Nóvoa, A., Bobola, N. and Mallo, M. (2005). Hox genes specify vertebral types in the presomitic mesoderm. *Genes Dev.* **19**, 2116–2121.
- Casaca, A., Nóvoa, A. and Mallo, M. (2016). *Hoxb6* can interfere with somitogenesis in the posterior embryo through a mechanism independent of its rib-promoting activity. *Development* **143**, 437–448.
- Condie, B. G. and Capecchi, M. R. (1994). Mice with targeted disruptions in the paralogous genes *hoxa-3* and *hoxd-3* reveal synergistic interactions. *Nature* **370**, 304–307.
- Dehni, G., Liu, Y., Husain, J. and Stifani, S. (1995). TLE expression correlates with mouse embryonic segmentation, neurogenesis, and epithelial determination. *Mech. Dev.* **53**, 369–381.
- Deries, M. and Thorsteinsdóttir, S. (2016). Axial and limb muscle development: dialogue with the neighbourhood. *Cell. Mol. Life Sci.* **73**, 4415–4431.
- Derynck, R., Zhang, Y. and Feng, X.-H. (1998). Transcriptional Activators of TGF- $\beta$  Responses: Smads. *Cell* **95**, 737–740.
- Deschamps, J., van den Akker, E., Forlani, S., De Graaff, W., Oosterveen, T., Roelen, B. and Roelfsema, J. (1999). Initiation, establishment and maintenance of Hox gene expression patterns in the mouse. *Int. J. Dev. Biol.* **43**, 635–650.
- Duboule, D. (1998). Vertebrate hox gene regulation: clustering and/or colinearity? *Curr. Opin. Genet.*

- Dev.* **8**, 514–518.
- Duboule, D.** (2007). The rise and fall of Hox gene clusters. *Development* **134**, 2549–60.
- Duboule, D. and Dollé, P.** (1989). The structural and functional organization of the murine HOX gene family resembles that of Drosophila homeotic genes. *EMBO J.* **8**, 1497–505.
- Duboule, D. and Morata, G.** (1994). Colinearity and functional hierarchy among genes of the homeotic complexes. *Trends Genet.* **10**, 358–364.
- Dubrulle, J. and Pourquié, O.** (2004). Coupling segmentation to axis formation. *Development* **131**, 5783–5793.
- Dubrulle, J., McGrew, M. J. and Pourquié, O.** (2001). FGF Signaling Controls Somite Boundary Position and Regulates Segmentation Clock Control of Spatiotemporal Hox Gene Activation. *Cell* **106**, 219–232.
- Economides, K. D., Zeltser, L. and Capecchi, M. R.** (2003). Hoxb13 mutations cause overgrowth of caudal spinal cord and tail vertebrae. *Dev. Biol.* **256**, 317–330.
- Favier, B. and Dollé, P.** (1997). Developmental functions of mammalian Hox genes. *Mol. Hum. Reprod.* **3**, 115–131.
- Forlani, S., Lawson, K. A. and Deschamps, J.** (2003). Acquisition of Hox codes during gastrulation and axial elongation in the mouse embryo. *Development* **130**,.
- Fortini, M. E., Artavanis-Tsakonas, S., Artavanis-Tsakonas, S., Simpson, P., Bang, A. G., Posakony, J. W., Busseau, I., Diederich, R. J., Xu, T., Artavanis-Tsakonas, S., et al.** (1993). Notch: Neurogenesis is only part of the picture. *Cell* **75**, 1245–1247.
- Galant, R. and Carroll, S. B.** (2002). Evolution of a transcriptional repression domain in an insect Hox protein. *Nature* **415**, 910–913.
- Garcia-Fernández, J.** (2005). Hox, ParaHox, ProtoHox: facts and guesses. *Heredity (Edinb.)* **94**, 145–152.
- Gasperowicz, M., Surmann-Schmitt, C., Hamada, Y., Otto, F. and Cross, J. C.** (2013). The transcriptional co-repressor TLE3 regulates development of trophoblast giant cells lining maternal blood spaces in the mouse placenta. *Dev. Biol.* **382**, 1–14.
- Gilbert, S.F.** (2013) *Developmental Biology* Sunderland, MA, Sinauer Associates Inc.
- Goetz, S. C., Ocbina, P. J. R. and Anderson, K. V.** (2009). The primary cilium as a Hedgehog signal transduction machine. *Methods Cell Biol.* **94**, 199–222.
- Gray, P. A., Fu, H., Luo, P., Zhao, Q., Yu, J., Ferrari, A., Tenzen, T., Yuk, D.-I., Tsung, E. F., Cai, Z., et al.** (2004). Mouse Brain Organization Revealed Through Direct Genome-Scale TF Expression Analysis. *Science (80-. )*. **306**, 2255–2257.
- Guerreiro, I., Casaca, A., Nunes, A., Monteiro, S., Nóvoa, A., Ferreira, R. B., Bom, J. and Mallo, M.** (2012). Regulatory role for a conserved motif adjacent to the homeodomain of Hox10 proteins. *Development* **139**, 2703–2710.
- Heldin, C.-H., Miyazono, K. and ten Dijke, P.** (1997). TGF- $\beta$  signalling from cell membrane to nucleus through SMAD proteins. *Nature* **390**, 465–471.

- Horan, G. S., Ramírez-Solis, R., Featherstone, M. S., Wolgemuth, D. J., Bradley, A. and Behringer, R. R.** (1995). Compound mutants for the paralogous *hoxa-4*, *hoxb-4*, and *hoxd-4* genes show more complete homeotic transformations and a dose-dependent increase in the number of vertebrae transformed. *Genes Dev.* **9**, 1667–77.
- Jennings, B. H., Pickles, L. M., Wainwright, S. M., Roe, S. M., Pearl, L. H. and Ish-Horowicz, D.** (2006). Molecular Recognition of Transcriptional Repressor Motifs by the WD Domain of the Groucho/TLE Corepressor. *Mol. Cell* **22**, 645–655.
- Jennings, B. H., Ish-Horowicz, D., Stifani, S., Blaumueller, C., Redhead, N., Hill, R., Artavanis-Tsakonas, S., Hartley, D., Preiss, A., Artavanis-Tsakonas, S., et al.** (2008). The Groucho/TLE/Grg family of transcriptional co-repressors. *Genome Biol.* **9**, 205.
- Jime, G., Paroush, Z. and Ish-horowicz, D.** (1997). Groucho acts as a corepressor for a subset of negative regulators, including Hairy and Engrailed. *Genes Dev.* **11**: 3072–3082.
- Ju, B.-G., Solum, D., Song, E. J., Lee, K.-J., Rose, D. W., Glass, C. K., Rosenfeld, M. G., Lee, S. C., Nakatani, Y., Shi, Y., et al.** (2004). Activating the PARP-1 sensor component of the groucho/TLE1 corepressor complex mediates a CaMKinase IIdelta-dependent neurogenic gene activation pathway. *Cell* **119**, 815–29.
- Kageyama, R., Ohtsuka, T., Hatakeyama, J. and Ohsawa, R.** (2005). Roles of bHLH genes in neural stem cell differentiation. *Exp. Cell Res.* **306**, 343–348.
- Kessel, M. and Gruss, P.** (1991). Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. *Cell* **67**, 89–104.
- Leon, C. and Lobe, C. G.** (1997). Grg3, a murine groucho-related gene, is expressed in the developing nervous system and in mesenchyme-induced epithelial structures. *Dev. Dyn.* **208**, 11–24.
- Lewis, E. B.** (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565–570.
- Li, D. and Roberts, R.** (2001). Human Genome and Diseases: WD-repeat proteins: structure characteristics, biological function, and their involvement in human diseases. *Cell. Mol. Life Sci.* **58**, 2085–2097.
- Mallo, M., Wellik, D. M. and Deschamps, J.** (2010). Hox genes and regional patterning of the vertebrate body plan. *Dev. Biol.* **344**, 7–15.
- McGinnis, W. and Krumlauf, R.** (1992). Homeobox genes and axial patterning. *Cell* **68**, 283–302.
- Neer, E. J., Schmidt, C. J., Nambudripad, R. and Smith, T. F.** (1994). The ancient regulatory-protein family of WD-repeat proteins. *Nature* **371**, 297–300.
- Noyes, M. B., Christensen, R. G., Wakabayashi, A., Stormo, G. D., Brodsky, M. H. and Wolfe, S. A.** (2008). Analysis of Homeodomain Specificities Allows the Family-wide Prediction of Preferred Recognition Sites. *Cell* **133**, 1277–1289.
- Ocbina, P. J. R., Eggenschwiler, J. T., Moskowitz, I. and Anderson, K. V** (2011). Complex interactions between genes controlling trafficking in primary cilia. *Nat Genet* **43**, 547–553.
- Paroush, Z., Finley, R. L., Kidd, T., Wainwright, S. M., Ingham, P. W., Brent, R. and Ish-Horowicz, D.** (1994). Groucho is required for *Drosophila* neurogenesis, segmentation, and sex determination and interacts directly with hairy-related bHLH proteins. *Cell* **79**, 805–15.

- Pearson, J. C., Lemons, D. and McGinnis, W.** (2005). Modulating Hox gene functions during animal body patterning. *Nat. Rev. Genet.* **6**, 893–904.
- Pedersen, L. B., Veland, I. R., Schröder, J. M. and Christensen, S. T.** (2008). Assembly of primary cilia. *Dev. Dyn.* **237**, 1993–2006.
- Pourquié, O.** (2003). The Segmentation Clock: Converting Embryonic Time into Spatial Pattern. *Science*. **301**, 328–330.
- Resende, T. P., Ferreira, M., Teillet, M.-A., Tavares, A. T., Andrade, R. P. and Palmeirim, I.** (2010). Sonic hedgehog in temporal control of somite formation. *Proc. Natl. Acad. Sci.* **107**, 12907–12912.
- Ronshaugen, M., McGinnis, N. and McGinnis, W.** (2002). Hox protein mutation and macroevolution of the insect body plan. *Nature* **415**, 914–917.
- Savage, C., Das, P., Finelli, A. L., Townsend, S. R., Sunt, C.-Y., Bairdt, S. E. and Padgett, R. W.** (1996). *Caenorhabditis elegans* genes *sma-2*, *sma-3*, and *sma-4* define a conserved family of transforming growth factor 3 pathway components (signal transduction/pattern formation/bone morphogenetic protein/multigene family). *Dev. Biol.* **93**, 790–794.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al.** (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682.
- Sekelsky, J. J., Newfeld, S. J., Raftery, L. A., Chartoff, E. H. and Gelbart, W. M.** (1995). Genetic characterization and cloning of mothers against dpp, a gene required for decapentaplegic function in *Drosophila melanogaster*. *Genetics* **139**, 1347–58.
- Singh, B. N., Koyano-Nakagawa, N., Donaldson, A., Weaver, C. V, Garry, M. G. and Garry, D. J.** (2015). Hedgehog Signaling during Appendage Development and Regeneration. *Genes (Basel)*. **6**, 417–35.
- Sirard, C., de la Pompa, J. L., Elia, A., Itie, A., Mirtsos, C., Cheung, A., Hahn, S., Wakeham, A., Schwartz, L., Kern, S. E., et al.** (1998). The tumor suppressor gene *Smad4/Dpc4* is required for gastrulation and later for anterior development of the mouse embryo. *Genes Dev.* **12**, 107–19.
- Smith, T. F., Gaitatzes, C., Saxena, K. and Neer, E. J.** (1999). The WD repeat: a common architecture for diverse functions. *Trends Biochem. Sci.* **24**, 181–5.
- Srour, M. K., Fogel, J. L., Yamaguchi, K. T., Montgomery, A. P., Izuhara, A. K., Misakian, A. L., Lam, S., Lakeland, D. L., Urata, M. M., Lee, J. S., et al.** (2015). Natural Large-Scale Regeneration of Rib Cartilage in a Mouse Model. *J. Bone Miner. Res.* **30**, 297–308.
- St-Jacques, B., Hammerschmidt, M. and McMahon, A. P.** (1999). Indian hedgehog signaling regulates proliferation and differentiation of chondrocytes and is essential for bone formation. *Genes Dev.* **13**, 2072–2086.
- Stirnimann, C. U., Petsalaki, E., Russell, R. B. and Müller, C. W.** (2010). WD40 proteins propel cellular networks. *Trends Biochem. Sci.* **35**, 565–574.
- Sudo, H., Takahashi, Y., Tonegawa, A., Arase, Y., Aoyama, H., Mizutani-Koseki, Y., Moriya, H., Wilting, J., Christ, B. and Koseki, H.** (2001). Inductive Signals from the Somatopleure Mediated by Bone Morphogenetic Proteins Are Essential for the Formation of the Sternal Component of



- Avian Ribs. *Dev. Biol.* **232**, 284–300.
- Tepass, U., Hartenstein, V., Azpiazu, N., Frasch, M., Bellen, H. J., O’Kane, C. J., Wilson, D., Grossniklaus, U., Pearson, R. K., Gehring, W. J., et al.** (1995). Neurogenic and proneural genes control cell fate specification in the *Drosophila* endoderm. *Development* **121**, 393–405.
- Van Hateren, N., Belsham, A., Randall, V. and Borycki, A. G.** (2005). Expression of avian Groucho-related genes (Grgs) during embryonic development. *Gene Expr. Patterns* **5**, 817–823.
- Vinagre, T., Moncaut, N., Carapuço, M., Nóvoa, A., Bom, J. and Mallo, M.** (2010). Evidence for a Myotomal Hox/Myf Cascade Governing Nonautonomous Control of Rib Specification within Global Vertebral Domains. *Dev. Cell* **18**, 655–661.
- Wellik, D. M.** (2007). Hox patterning of the vertebrate axial skeleton. *Dev. Dyn.* **236**, 2454–2463.
- Wellik, D. M. and Capecchi, M. R.** (2003). Hox10 and Hox11 genes are required to globally pattern the mammalian skeleton. *Science* **301**, 363–7.
- Woltering, J. M. and Durston, A. J.** (2006). The zebrafish hoxDb cluster has been reduced to a single microRNA. *Nat. Genet.* **38**, 601–602.
- Yang, H., Wang, H., Shivalila, C. S., Cheng, A. W., Shi, L. and Jaenisch, R.** (2013). One-step generation of mice carrying reporter and conditional alleles by CRISPR/cas-mediated genome engineering. *Cell* **154**, 1370–1379.
- Young, T., Rowland, J. E., van de Ven, C., Bialecka, M., Novoa, A., Carapuco, M., van Nes, J., de Graaff, W., Duluc, I., Freund, J.-N., et al.** (2009). Cdx and Hox Genes Differentially Regulate Posterior Axial Growth in Mammalian Embryos. *Dev. Cell* **17**, 516–526.
- Zákány, J., Kmita, M., Alarcon, P., de la Pompa, J.-L. and Duboule, D.** (2001). Localized and Transient Transcription of Hox Genes Suggests a Link between Patterning and the Segmentation Clock. *Cell* **106**, 207–217.
- Zhao, Y. and Potter, S. S.** (2001). Functional specificity of the Hoxa13 homeobox. *Development* **128**, 3197–3207.
- Zhao, Y. and Potter, S. S.** (2002). Functional Comparison of the Hoxa 4, Hoxa 10, and Hoxa 11 Homeoboxes. *Dev. Biol.* **244**, 21–36.



# **Appendix I: Standard Molecular Procedures**

### **Agarose gel analysis**

Most work involving DNA and RNA molecules used agarose gel electrophoresis for analysis and preparation. The standard protocol used agarose concentrations of 2% or 0,8% (dissolved in 1X TAE buffer) and containing 2.5 µl of RedSafe® (1X), a nucleic acid staining solution, per 50 ml of agarose. Before loading the gel, 6X loading dye (Appendix II) was added to each sample (1X final concentration).

### **Phenol/Chloroform DNA purification**

For cloning the PCR products, they were purified using the Phenol/Chloroform (P/C) technique. The volume of the reaction was taken to 100 µl with TE pH 8.0 and the same volume of phenol/chloroform was added to the samples and vortexed. The samples were then centrifuged for 4 minutes at RT, maximum speed. The aqueous phase was retrieved and the DNA precipitated by adding 1:10 volume of 3M NaOAc, pH 5.3 and 2.5 volumes of ethanol. This mixture was incubated for 1 hour on dry ice and the precipitate recovered by centrifugation for 30 minutes, 14000 rpm, at 4°C. Afterwards, the supernatant was discarded, the pellet air-dried and resuspended in 25 µl of Milli Q water. This DNA was ready for digestion.

### **DNA digestion and purification for subcloning**

DNA to be subcloned, either from PCR products or from a plasmid was digested with the appropriate enzymes. At the same time, 4 µg of the relevant vector (pBluescript KS (-), pDLL plasmids for microinjection or pCMV-Sport6.1 for cell transfection) were also digested with the appropriate enzymes. The relevant bands were run on a 1X TAE agarose gel (1%) containing Ethidium Bromide (0.2 µg/ml), the bands cut out of the gel and purified using the QIAEX II Gel Extraction kit according to the manufacturer's protocol. The bands were resuspended in 20 µl TE pH 8.0.

### **Insert-Vector Ligation**

Ligations were done using 50 ng of vector DNA, a 1 to 3 molar ratio of insert, 2 µl of 10x ligation buffer, 1 µl of T4 ligase and water up to 20 µl. A control reaction containing only the vector DNA was also included. The reactions were incubated for 1 hour at RT and then 5 µl were used for transformation.

When the TOPO protocol was used to clone PCR products, it was done following the manufacturer's specifications. Essentially the ligation reaction contained 4 µl of fresh PCR product, 1 µl of salt solution and 1 µl of TOPO vector. The reaction was incubated for 5 minutes at RT. This product was used for transformation.

### **Transformation**

Transformation was performed on DH5-α competent cells using the heat shock protocol. 50 µl of these cells were thawed on ice and incubated with the DNA for 20 minutes on ice. The mix was incubated at 42°C for 35 seconds and then 5 minutes on ice. 800 µl of LB medium (Appendix II) were then added and, the cells incubated at 37°C for 45 minutes with shaking. The culture was then centrifuged at 4000 rpm for 4 minutes and most of the supernatant removed. Cells were resuspended in 500 µl of LB medium and 50 µl of this was plated on LB agar plates with ampicillin (50 µg/ml) and incubated ON at 37°C.

### **Colony screening by PCR and digestion analysis**

Individual colonies were picked from the LB/amp plates using pipette tips and grown in 200 µl of LB medium with ampicillin (50 µg/ml) at 37°C with shaking for three hours. Then, a PCR reaction was performed using primers specific for the insert, using 5 µl of the culture as template.

To confirm the results, those colonies that were positive in the PCR screen were further grown in 3 ml LB medium with ampicillin (50 µg/ml) at 37°C with shaking ON. 1 ml of the grown cultures was centrifuged at 4000 rpm for four minutes and the pellet resuspended in 100 µl of TE containing RNase (10 µg/ml). 300 µl of TENS (Appendix II) were added and vortexed until sticky. Genomic DNA was precipitated by adding 150 µl of 3M potassium acetate (pH 5.2), vortexed and this mixture centrifuged for 5 minutes at 14000 rpm. The supernatant was transferred into a new tube containing 800 µl of 100% ethanol, vortexed and centrifuged for five minutes at 14000 rpm. Finally, the supernatant was removed, the pellet air-dried and resuspended in 50 µl of TE.

5 µl of the isolated plasmid DNA were used in several digestions, using restriction enzymes that allowed to confirm insert presence and its orientation. Positive colonies were further grown and used for the relevant procedures

### **Plasmid DNA midiprep**

To produce large amounts of pure DNA, the relevant clones were picked into 100 ml of LB medium with ampicillin (50 µg/mL) and incubated at 37°C with shaking ON. The purification was done using the NucleoBond® plasmid DNA purification kit as specified in the manufacturer's protocol. The pellet was resuspended in 100 µl of TE pH 8.0 and DNA concentration was measured using NanoDrop 1000 Spectrophotometer.

### **Sanger sequencing**

This step was necessary to confirm the sequence of the different constructs amplified by PCR. Primers used for sequencing reactions flanked the multiple cloning site (M13 Reverse and M13 Forward).

DNA sequencing was performed following the Sanger method, using the BigDye® Terminator v1.1 Cycle Sequencing Kit. The reaction amounts were as follows: 2 µl of BigDye® terminator sequencing buffer (5X), 2 µl BigDye® terminator ready reaction mix, 500 ng of template DNA, 1 µl of primer and H<sub>2</sub>O up to 10 µl. Cycles were run in a PCR machine according to the conditions shown in Appendix II. Samples were then precipitated by adding H<sub>2</sub>O (10 µl), NaOAc 3M (2 µl) and EtOH 95% (50 µl) and incubated for 30 minutes at RT. Afterwards, this mix was centrifuged for 30 minutes (14000 rpm) at 4°C and the supernatant discarded. The pellet was rinsed with 250 µl of EtOH 70% and centrifuged for 15 minutes (14000 rpm) at 4°C. The supernatant was discarded and the dried pellet was sent to the Sequencing Unit of the IGC. The resulting sequences were analysed using SnapGene and BLAST.

# **Appendix II – Buffers, Media and Other Solutions**

## Common Buffers

1X TE	
EDTA	1 mM
Tris-HCl	10 mM

1X TAE	
EDTA pH 8	1 mM
Acetic Acid	20 mM
Tris Base	40 mM

TENS Buffer	
Tris pH 7.5	10 mM
EDTA	1 mM
NaOH	0.1 M
SDS	0.5%

## Ethidium Bromide Gel

1% Gels	
Agarose	1%
Ethidium Bromide	0.2 µg/ml
1x TAE	Up to final volume

6X Loading Dye	
Glycerol	30%
Bromophenol blue	0.25%

Molecular Ladders	
DNA Ladders	Thermo Scientific™ GeneRuler™ 1kb (MAN0013004) and NZYDNA Ladder VI (MB08901)
Protein Ladder	NZYColour Protein Marker II (MB09002)

## Bacterial Growth and Plasmid Purification

Lysogeny Broth (LB) medium	
Tryptone	1%
Yeast extract	0.5%
NaCl	1%

<b>DNA purification Kits</b>	
<b>NZYGelpure</b>	NZYTech (MB01101)
<b>QIAEX II® Gel Extraction Kit</b>	QIAGEN (#20051)
<b>NZYMiniprep</b>	NZYTech (MB01001)
<b>NucleoBond® Xtra Midi</b>	Macherey-Nagel (740410.10)

## Genotyping

<b>Tail Lysis Buffer</b>	
<b>Tris-HCl pH 8.3</b>	10 mM
<b>MgCl<sub>2</sub></b>	2.5 mM
<b>KCl</b>	50 mM
<b>Tween-20</b>	0.45%
<b>Nonidet P40 (NP40)</b>	0.45%
<b>Gelatin</b>	0.1 mg/ml

<b>Laird's Buffer</b>	
<b>Tris-HCl, pH 8.5</b>	100 mM
<b>EDTA</b>	5 mM
<b>SDS</b>	0.2%
<b>NaCl</b>	200 mM

<b>Yolk Sac Lysis Buffer</b>	
<b>Tris-HCl pH 8.3</b>	10 mM
<b>MgCl<sub>2</sub></b>	2 mM
<b>KCl</b>	50 mM
<b>Tween-20</b>	0.45%
<b>NP40</b>	0.45%

## PCR Conditions for Genotyping and Cloning

<b>Enzymes</b>	
<b>Taq DNA polymerase</b>	Thermo Scientific (EP0281) and NZYTech (MB00101)
<b>Pfu DNA polymerase</b>	Thermo Scientific (EP0572)
<b>Reverse Transcriptase</b>	NZYTech (MB12401)



HoxA10/ Hoxa10ΔC1/Hoxa10ΔC1p1/ Hoxa10ΔC1p2	
DNA	1 μl
10x Buffer	2.5 μl (1x)
DMSO	2 μl (8% v/v)
Forward Primer (25mM)	0.25 μl
Reverse Primer (25mM)	0.25 μl
dNTPs (25mM)	0.2 μl
Taq (5 U/μl)	0.2 μl
H <sub>2</sub> O	Up to 25 μl

Temperature	Time	35x
95°C	4 min	
95°C	45 sec	
65°C	1 min	
72°C	45 sec	
72°C	10 min	

Grg3/IFT144 Cloning	
DNA	1 μl
10x Buffer	2.5 μl (1x)
Forward Primer (25mM)	0.25 μl
Reverse Primer (25mM)	0.25 μl
dNTPs (25mM)	0.2 μl
Taq (5 U/μl)	0.2 μl
H <sub>2</sub> O	Up to 25 μl

Temperature	Time	35x
95°C	4 min	
95°C	45 sec	
60°C	1 min	
72°C	45 sec	
72°C	10 min	

#### Sanger Sequencing PCR conditions

Sanger Sequencing	
DNA	500 ng
BigDye® terminator sequencing buffer (5X)	2 μl
BigDye® terminator ready reaction mix	2 μl
Sequencing Primer (3,2 pmol/μ)	1 μl
H <sub>2</sub> O	Up to 10 μl

Temperature	Time	25x
96°C	1 min	
96°C	10 sec	
50°C	5 sec	
60°C	4 min	

#### Skeletal Staining

Alcian Blue Solution	
Alcian Blue 8 GX	150 mg/L
Ethanol	80%
Acetic acid	20%

Alizarin Red Solution	
Alizarin Red S	50 mg/L
KOH	2%

## Western Blot

Tris Glycine	
Tris Base	25 mM
Glycine	192 mM

Running Buffer	
Tris Glycine	1X
10% SDS	0.1%

Western Blot Transfer Buffer	
10X Tris Glycine	1X
Methanol	20%

Western Blot 6x Loading Buffer	
Tris-HCl pH 6.8	100 mM
SDS	4%
Glycerol	20%
$\beta$ -mercaptoethanol	5%
Bromophenol blue	0.02%

10%-12% SDS Polyacrylamide Resolving Gel	
30%acrylamide/bisacrylamide	10%-12%
1.5M Tris-HCl pH 8.8	390 mM
10% SDS	0.1%
10% ammonium persulfate (APS)	0.1%
Tetramethylethylenediamine (TEMED)	0.04%

5% SDS Polyacrylamide Stacking Gel	
30%acrylamide/bisacrylamide	5%
1M Tris-HCl pH 6.8	125 mM
10% SDS	0.1%
10% Ammonium Persulfate (APS)	0.1%
TEMED	0.04%

PBT	
10% Tween 20	1%
Dulbecco's Phosphate Buffered Saline (PBS)	Up to final volume

<b>5% Blocking Solution</b>	
<b>Powder Milk</b>	5%
<b>10% Tween-20</b>	0.1%
<b>Dulbecco's Phosphate Buffered Saline (PBS)</b>	Up to final volume

<b>Western Primary Antibodies</b>	
<b>Anti-FLAG</b>	1:1000; Sigma-Aldrich (F1804)
<b>Anti-c-MYC</b>	1:1000; Clontech (631206)
<b>Anti-IFT144</b>	1:750; Proteintech (13647-1-AP)

<b>Western Secondary Antibodies</b>	
<b>Anti-Mouse (DyLight® 680)</b>	1:2000; Invitrogen (SA5-10170)
<b>Anti-Rabbit (DyLight® 680)</b>	1:2000; Invitrogen (SA5-10042)

## Cell Culture

<b>Feeder Media</b>	
<b>Dulbecco's Modified Eagle's Medium (DMEM)</b>	Sigma-Aldrich (D5796)
<b>Fetal Bovine Serum (FBS)</b>	Sigma-Aldrich (F7524)
<b>Penicillin and Streptomycin</b>	Sigma-Aldrich (P0781)
<b>L-glutamine</b>	Sigma-Aldrich (G7513)

<b>Transfection Media</b>	
<b>DMEM</b>	Sigma-Aldrich (D5796)
<b>10% FBS</b>	Sigma-Aldrich (F7524)
<b>100x L-glutamine</b>	Sigma-Aldrich (G7513)

<b>Trypsinization</b>	
<b>Trypsin-EDTA (0.25%)</b>	Sigma-Aldrich (T3924)
<b>PBS</b>	Sigma-Aldrich (D1408)

<b>Transfection</b>	
<b>Lipofectamin™ 2000 Transfection Reagent</b>	Invitrogen™ (11668019)

## Co-IP

Washing Buffer	
Tris-HCl pH 7.5	50 mM
NaCl	150 mM
MgCl <sub>2</sub>	6 mM
IgePal	1%
Sodium Deoxycholate	0.5%
Protease Inhibitors	1X

Washing Buffer	
Hepes pH 7.4	10 mM
NaCl	150 mM
IgePal	0.02%
Protease Inhibitors	1X

Beads	
Dynabeads® Protein G	Invitrogen (10003D)
Anti-FLAG® M2 Magnetic Beads	Sigma-Aldrich (M8823)

## Immunocytochemistry

PBS-T	
Triton	0.3%
PBS	Sigma-Aldrich (D1408)

PBS-FBS	
FBS	0.5% or 10%
PBS	Sigma-Aldrich (D1408)

Primary Antibodies	
Anti-FLAG	1:500; Sigma-Aldrich (F1804)
Anti-c-MYC	1:500; Clontech (631206)
Anti-IFT144	1:200; Proteintech (13647-1-AP)

Secondary Antibodies	
Anti-Mouse Alexa Fluor® 568	1:500; Abcam (ab175473)
Anti-Rabbit (Alexa Fluor® 488)	1:500; Abcam (ab150077)

Mounting Medium	
VECTASHIELD® Mounting Medium	Vector Laboratories (H-1000)

### *In situ* Hybridization

Hybridization Solution 1	
Formamide	50%
Saline-Sodium Citrate (SSC) pH 5	5X
EDTA pH 8	5 mM
10 % Tween-20	2%
Heparine	0.1 mg/ml
Yeast tRNA	0.05 mg/ml

Hybridization Solution 2	
Formamide	50%
Saline-Sodium Citrate (SSC) pH 5	5X
EDTA pH 8	5 mM
10 % Tween-20	2%

TBST	
Tris-Buffered Saline (TBS)	1X
10 % Tween-20	1%

MABT	
Maleic Acid Buffer	1X
10 % Tween-20	1%

Blocking Solution (1%-10%)	
Blocking Reagent	1%; Roche (000000011096176001)
Sheep Serum	1% or 10%
MABT	Up to final volume

NTMT	
Tris-HCl pH 9.5	100 mM
NaCl	100 mM
MgCl <sub>2</sub>	50 mM
10% Tween-20	10%

Antibody	
Anti-Digoxigenin-AP	Roche (000000011093274910)

Developing Solution	
NBT-BCIP® Solution	2%; Sigma-Aldrich (72091)
NTMT	Up to final volume

## **Appendix III: Sequences and Primers**

Neither the c-MYC-tag or FLAG-tag are shown in any of the following DNA sequences.

**Grg3 (WD40 domain cloned into pCMV Sport6.1):**

CGATGGGCAGATGCAACCTGTGCCCTTCCCCCATGATGCACTAGCAGGCCCTGGCATTCC  
CAGGCATGCCCCGGCAGATCAATACGCTCAGCCATGGAGAGGTGGTATGTGCTGTGACCAT  
CAGCAACCCACACGACACGTCTACACAGGCGGCAAGGGCTGTGTGAAGATATGGGACA  
TCAGCCAGCCGGGCAGCAAGAGTCCCATCTCCAGCTGGACTGCCTGAACAGGGACAAC  
TACATCCGCTCGTGCAAGCTTCTCCCCGATGGGCGCACGCTCATTGTGGGTGGTGAGGCC  
AGCACGCTCACCATCTGGGACCTGGCCTCACCCACACCCCGCATCAAGGCTGAGCTGACG  
TCCTCGGCTCCAGCCTGTTATGCCCTGGCCATCAGTCCTGATGCCAAAGTCTGTTTTCT  
GCTGCAGCGACGGGAACATTGCGGTTTGGGATCTGCACAACCAGACCCTGGTCAGGCAGT  
TCCAGGGCCACACAGATGGGGCCAGCTGTATAGACATCTCTCATGATGGCACTAAGCTGT  
GGACCGGGGGCCTGGACAACACCGTGCGCTCCTGGGACCTACGTGAAGGACGGCAGTTA  
CAGCAACACGATTTACCTCCCAGATCTTCTCCCTGGGTTACTGCCCCACTGGGGAGTGG  
CTGGCCGTGGGCATGGAGAGCAGCAATGTGGAGGTCCTGCACCACACTAAGCCCGACAA  
ATACCAGCTGCACCTGCACGAGAGCTGCGTGCTGTCCCTCAAGTTCGCCTATTGTGGCAA  
GTGGTTTGTGAGCACTGGGAAAGACAACCTTCTCAATGCCTGGAGGACGCCTTATGGAGC  
CAGCATCTTCCAGTCAAAAGAATCCTCATCTGTCTTGAGCTGTGACATTTACGCGGATGA  
CAAATATATTGTAACAGGCTCTGGTGACAAGAAGGCCACAGTTTACGAGGTCATCTACTG  
AACAAGGACTCTAACAGGCCTGTCAAACCTCTGGGAGAGACACCCACGTGGCCCGCGGCC

**Grg3 KO Result**

The sequence that is replaced is in bold and the inserted stop codons are underlined.

TGTGCCCACTCCTCTCCCCTATGTTGT**CCTTGTGGGCTCCAGGCACCCCATCAACCCGG**  
**GCAGCCGGGATTTAAATTC**ACTGTGGCCGAGTCCTGTTAGTAGTAGGACAGGATCA  
AAGACGAATTCAGTT**CCTGCAAGCTCAGTATCACAGGTAAGGCGGGTGGGGGGTG**  
**GCCCCGGGCTGGCGAGGAGGGCGGCCTT**

**IFT144 (cloned in pCMV Sport6.1)**

GGATCCATGAAGCGTGTTTTCTCCCTGCTAGAAAAGTCTTGGCTTGGTGCTCCGATACAAT  
TTGCCTGGCAAAAATCATCAGGAAACTACCTTGCAGTAACAGGAGCTGATTATATTGTTA  
AAATCTTTGATCGCCATGGCCAAAAAAGAAGTGAAATTAGCTTGCCTGGCAACTGTGTTA  
CCATGGATTGGGATAAAGATGGCGATATCCTGGCAGTGATTGCTGAGAAGTCTAGTTGCA  
TTTATCTATGGGATGCCAACACAAATAAAACCAGCCAGCTGGACAATGGCATGAGGGAT  
CAAATGTCTTTCTTCTTTGGTCAAAAATTGGAAGTTTCCTGGCTGTTGGGACCATTAAAG  
GAAATTTGCTCATTTATAATCATCAGACATCTCGAAAGATTCTGTTCTTGGAAAACATAC  
TAAGAAAATCACATGTGGATGTTGGAATTCAGAGAATCTCCTTGCTTTGGGAGGTGAAGA  
TAAAATGATTACAGTTAGTAACCAGGAAGGCGACACAATAAGACAGACCCCAAGTGAAAT  
CAGAGCCAAGCGACATCAAGTTCTCCATGAGCAAGACAGATGAGCGAATTTCTTCTGCTG  
AGAACACAATAAGTGCAGTGGTTGGCAAGAAAATGCTGTTTCTTTTTTCATCTGAATGAAC  
CAGATAACCCGGTGGATCTGGAGTTTCAGCAAGCCTATGGCAACATTGTCTGCTATAGTT  
GGTATGGAGATGGCTACATCATGATTGGCTTTTCCCGAGGGACGTTTTTGGCTATTTCTAC  
TCACTTTCCGGAAGTTGGGCAAGAGATATTTAAGGCTCGTGACCATAAGGATAATCTAAC  
CAGTGTGGCCTTGTCACAGACTCTGAACAAAGCTGCCACATGTGGCGATAACTGCATAAA  
AATCCATGATCTGACAGAATTGAGAGACATGTATGCTATAATTAATCTGGATGATGAGAA  
TAAAGGGCTGGGTACCTTATCCTGGACTGATGATGGTCAGTTGCTAGCACTGTCTACCCA  
AAGAGGCTCACTGCATGTCTTCTGACCAAGTTGCCCATCCTCGGGGACGCCTGTACAC

AAGGATTGCGTATCTCACCTCCCTCCTTGAGGTCACCGTGGCCAACCTCATTGAAGGAGA  
GCCGCCAATCACAGTCTCTGTGGATGTGGAACCCACCTTTGTCGCAGTAGGGCTCTATCA  
TCTGGCCGTGGGGATGAATAACCGGGCTTGGTTTTATGTCCTTGGTGAAAATGTTGTCAA  
AAAGTTAAAAGATGTGGAATATCTGGGAACCGTGGCCAGCATCTGCCTTCATTCTGACTA  
CGCCGCTGCACTCTTTGAAGGCCAAAATCCAGTTACATTTGATAGAAAATGAAATGTTGGA  
CGCTCAGGAAGAGCGTGAGACTCGGCTCTTTCCAGCAGTGGATGATAAGTGCCGGATTTT  
ATGCCACGCCCTAACTAGTGATTTCCCTCATCTACGGAACCTGATACTGGCATCATTCACTAT  
TTCTTCATCGAAGACTGGCAGTTCGTTAATGATTACCGGCATCCTGTTGGTGTGAAGAAG  
CTATTTTCTGATCCAAATGGAACCAGATTGGTTTTTCATTGATGAGAAGAGTGATGGATTT  
GTTTACTGTCCTGTTAATGATGCGACCTATGAGATTCCAGACTTCTCACCAACCATTAAAG  
GTGTTCTTTGGGAAAACCTGGCCGATGGACAAAGGTGTCTTTATCGCCTATGATGATGACA  
AGGTGTATACATATGCGTTTCACAAGGACACCATCCAAGGATCCAAGGTTATTTTGGCTG  
GCAGCACCAAACCTTCCCTTCTCCATAAGCCTTTGCTGTTATACAATGGAGAACTGACCTG  
CCAGACACAGAGTGGGAAAATCAACTCCATCTACCTCAGCACCCACAGCTTCCTTGGCAG  
CATGAAAGACACGGAGCCTACTGACCTGAGGCAAATGCTGACGCAGACCCTGCTGCTCA  
AGCGGTTTTCTGATGCTTGGGATATATGCAAGATGCTAAATGACCGCACTTCGTGGAGTG  
AGCTGGCCAAAGCCTGTCTGCATCACATGGAGGTGGAGTTTGCTATCCGAGTGTCCCGGA  
CAATGGGGGATGTTGGCACAGTGATGTCGTTGGAACAAATAAAGGGAATCGAGGACTAC  
AATCTTTTGGCAGGACATCTCGCCATGTTTACTAATGACTTCAACCTGGCCCAGGACCTGT  
ACCTGGCATCCAACCTGCCCTGTGGCAGCCCTGGAGATGCGGCGGGACCTGCAGCACTGGG  
ACAGCGCTCTGCAGCTGGCAAAGCGCCTGGCCCCGGACCAGATACCCTTCATATCCAAAG  
AGTACGCCATCCAGCTGGAGTTCACAGGCGATTATGTAAACGCTCTGGCTCATTACGAGA  
AGGGCATCACCGGTGATAATAAGGAACACGACGAAGTGTGCCTGGCCGGAGTGGCTCAG  
ATGTCCATTTCGAATGGGGGACATCCGCAGAGGGGGCTAACCAAGCCCTCAAGCACCCCAG  
CAGGGTCTCAAAAGAGACTGTGGAGCCATTCTGGAGAACATGAAGCAATTTTCAGAAG  
CTGCCCAGCTGTACGAAAAGGGCCAATATTATGACAGAGCTGCCTCGGTCTACATCCGCT  
GCAAGAACTGGGCAAAAGTTGGCGAACTTCTCCCTCATGTCTCCTCTCCTAAGATCCACTT  
GCAGTATGCCAAAGCCAAGGAGGCAGACGGAAGGTACAAGGAAGCCGTGGTGGCGTATG  
AAAATGCAAAGCAATGGAACAGTGTATCCGCATCTACCTGGACCACCTCAACAACCCCCG  
AAAAGGCCGTGAGCATCGTCAGAGAGACCCAGTCTCTGGACGGAGCCAAGATGGTAGCC  
AGGTTCTTTCTGCAGCTTGGTGACTATGGGTCTGCCATCCAGTTTCTGGTTCTGTCCAAAT  
GTAACAATGAAGCCTTCACCTTGGCTCAGCAGCACAAACAAAATGGAAATCTACGCAGAC  
ATCATTGGTGCTGAAGACACAATAATGAAGACTATCAAAGTATCGCCTTATATTTTGAA  
GGAGAAAAAAGACATTTTCAGGCTGGAAAATTCTTCTTACTGTGTGGCCAGTATTCACGG  
GCACTAAAGCACTTCCTGAAATGCCCAAGCTCAGAAGATAATGTGGCAATAGAAATGGC  
AATCGAAACTGTGGGCCAGGCCAAAGATGAACTGCTGACCAATCAGCTGATCGACCACC  
TGATGGGGGAGAGCGATGGCATGCCAAAGGACGCCAAGTACCTGTTCCGCTTGTACATG  
GCGCTAAAGCAGTACCGTGAAGCAGCCCGGACCGCCATCATCATCGCCAGAGAAGAGCA  
GTCTGCAGGAACTATCGGAATGCACACGATGTTCTTTTTCAGTATGTACGCAGAACTTAA  
AGCCCAGAAGATCAAGATCCCCTCCGAAATGGCCACCAACCTCATGATCCTGCACAGTTA  
CATTCTCGTGAAGATTCATGTTAAGAGTGGAGACCATATGAAGGGAGCGCGCATGCTCAT  
TCGGGTGGCCAAACATATCAGCAAGTTCCCATCACACATCGTGCCTATCCTGACGTCTAC  
TGTGATTGAGTGTATAGGGCAGGCCTGAAAACTCTGCCTTCAGCTTTGCAGCTATGCT  
GATGAGGCCTGAATACCGCAACAAAATTGATGCCAAGTACAAAAAGAAAATTGAGGCGA  
TGGTCAGGAGACCCGATACTTCAGAGACAGAAGAGGCCACCACCCCATGTCCATTCTGCC  
AGTTTCTTCTCCAGAATGTGAGCTCCTCTGTCCTGGCTGTAAAAACAACATTCCTATTG  
CATTGCAACAGGCCGACACATGTTGAAAGACGACTGGACAATGTGCCCCGATTGTGGCTT  
CCCTGCTCTGTACTCAGAATTCAAGATCTTACTAAACAGTGAAAGCACGTGTCCTATGTGT



TCAGAAAGATTAAACTCCAGTCAACTGAAAAAATTACAGACTGCTCGCAGTACCTACGG  
ACAGAGATGGAATAGTGAGCCGCAGGCCAGACAGTGCTCCCGAGTAGATGGCACCTCT  
CCTCTCGAG

#### Transgenic constructs used in this work:

##### Hoxa10

The C1 motif is shown in bold and underlined.

GGATCCTGCTCGGAGAGCCCTGCCGCGAACTCCTTTTTGGTCGACTCGCTCATCAGCTCAG  
GCAGAGGCGAGGCTGGTGTGGTGGCGGTAGCGCGGGGGGCGGTGGAGGTGGCTACTAC  
GCCACGGTGGGGTCTACCTGCCGCTGCCAGCGACCTGCCCT**TACGGGCTGCAAAGCTG**  
**CGGGCTCTTCCCCGCGCTGGGCAGCAAGCGTAAT**GAAGCGCCGTCGCCCCGAGGCGG  
TGGCGGTGGTGGCAGCGGGGGCCTGGGTCTGGGACGCATGGCTACGCGCCCCGCGCCCCCT  
AGACCTGTGGCTGGACGCGCCCCGCTCCTGCCGGATGGAGCCGCCCCGACGGGCCGCGC  
CACCAGAGCCACAACCCAGCAGCAGCAGCAGCAGCCGCGCCGCCCCCGCCGAGCCA  
CCTCAACCCAGCCACAGGCCACTTCGTGTTCTTTTGCAGAACATCAAAGAAGAGAGC  
TCCTACTGCCTCTACGATGCTGCGGACAAATGCCCCAAGGGCTCGGCCGCGCTGATCTG  
GCCCTTTCCCGCGGGGGCCCGCCCGCCGACGGCTGCGCCCTGGGCGCCTCCAGCGGAGTG  
CCAGTACCCGGCTACTTCCGCCTGTGCGAGGCCTACGGCACGGCCAAGGGCTTCGGCAGT  
GGCGGCGGCGGCACGCAGCAGCTCGCTAGTCCCTTTCTGCGCAGCCCCCGGGGCGCGGT  
TTCGACCCGCGCCCGCACTGGCCTCTGGCTCGACCGAGGCAGCCGGGAAGGAGCGAGT  
CCTAGACTCCACGCCACCACCACTCTGGTTTGCACCGGTGGCGGGCGGCTCGCAGGGCGA  
CGAGGAGGCACACGCGTCATCCTCGGCGGCTGAGGAGCTGTCTCCAGCCCCCTCAGAAAA  
CAGTAAAGCTTCGCCGGAGAAGGACTCCCTGGGCAGTTCCAAAGGCGAAAATGCAGCCA  
ACTGGCTCACAGCAAAGAGCGGGCCGGAAGAAACGCTGCCCTTACACGAAGCACCAGACG  
CTGGAGCTGGAGAAGGAGTTTCTATTCAACATGTACCTTACTCGAGAGCGGCGCCTAGAG  
ATCAGCCGTAGCGTCCACCTCACGGACAGACAAGTGAAAATCTGGTTTCAGAAATCGCAGG  
ATGAAACTGAAGAAAATGAACCGAGAAAACCGAATCCGGGAGCTCACAGCCAACCTTAA  
TTTTTCCTGA

##### Hoxa10ΔC1p1

What remains of the C1 motif is shown in bold and underlined.

GGATCCTGCTCGGAGAGCCCTGCCGCGAACTCCTTTTTGGTCGACTCGCTCATCAGCTCAG  
GCAGAGGCGAGGCTGGTGTGGTGGCGGTAGCGCGGGGGGCGGTGGAGGTGGCTACTAC  
GCCACGGTGGGGTCTACCTGCCGCTGCCAGCGACCTGCC**CCCCGCGCTGGGCAGCAA**  
**GCGTAAT**GAAGCGCCGTCGCCCCGAGGCGGTGGCGGTGGTGGCAGCGGGGGCCTGGGT  
CTGGGACGCATGGCTACGCGCCCCGCGCCCCCTAGACCTGTGGCTGGACGCGCCCCGCTCCT  
GCCGGATGGAGCCGCCCCGACGGGCCCGCCGCCACCGCAGCCACAACCCAGCAGCAGCAG  
CAGCAGCCGCGCCGCCCCCGCCGAGCCACCTCAACCCAGCCACAGGCCACTTCGTGT  
TCTTTTGCAGAACATCAAAGAAGAGAGCTCCTACTGCCTCTACGATGCTGCGGACAAA  
TGCCCCAAGGGCTCGGCCGCGCTGATCTGGCCCCCTTTCCCGCGGGGGCCCGCCCGCCGAC  
GGCTGCGCCCTGGGCGCCTCCAGCGGAGTGCCAGTACCCGGCTACTTCCGCCTGTGCGAG  
GCCTACGGCACGGCCAAGGGCTTCGGCAGTGGCGGCGGCGGCACGCAGCAGCTCGCTAG  
TCCCTTTCTGCGCAGCCCCCGGGGCGCGGTTTCGACCCGCGCCCGCACTGGCCTCTGG  
CTCGACCGAGGCAGCCGGGAAGGAGCGAGTCCTAGACTCCACGCCACCACCACTCTGG  
TTTGCACCGGTGGCGGCGGCTCGCAGGGCGACGAGGAGGCACACGCGTCATCCTCGGCG  
GCTGAGGAGCTGTCTCCAGCCCCCTCAGAAAACAGTAAAGCTTCGCCGGAGAAGGACTCC

CTGGGCAGTTCCAAAGGCGAAAATGCAGCCAACTGGCTCACAGCAAAGAGCGGCCGGAA  
GAAACGCTGCCCTTACACGAAGCACCAGACGCTGGAGCTGGAGAAGGAGTTTCTATTCA  
ACATGTACCTTACTCGAGAGCGGCGCCTAGAGATCAGCCGTAGCGTCCACCTCACGGACA  
GACAAGTGAAAATCTGGTTTCAGAATCGCAGGATGAACTGAAGAAAATGAACCGAGAA  
AACCGAATCCGGGAGCTCACAGCCAACTTTAATTTTCTCTGA

### **Hoxa10ΔC1p2**

GGATCCTGCTCGGAGAGCCCTGCCGCGAACTCCTTTTTGGTCGACTCGCTCATCAGCTCAG  
GCAGAGGCGAGGCTGGTGTGGTGGCGGTAGCGCGGGGGGCGGTGGAGGTGGCTACTAC  
GCCACGGTGGGGTCTACCTGCCGCTGCCAGCGACCTGCCCTACGGGCTGCAAAGCTG  
CGGGCTCTTCCCGTGCCCCGGAGGCGGTGGCGGTGGTGGCAGCGGGGGCCTGGGTCTG  
GGACGCATGGCTACGCGCCCGCGCCCCTAGACCTGTGGCTGGACGCGCCCCGCTCCTGCC  
GGATGGAGCCGCCGACGGGCCGCCGCCACCGCAGCCACAACCCAGCAGCAGCAGCAG  
CAGCCGCCGCCGCCGCCGCCGCCGCGCAGCCACCTCAACCCAGCCACAGGCCACTTCGTGTTCT  
TTTGCGCAGAACATCAAAGAAGAGAGCTCCTACTGCCTCTACGATGCTGCGGACAAATGC  
CCCAAGGGCTCGGCCGCCGCTGATCTGGCCCCCTTTCCCGCGGGGGCCCGCCGCCGACGGC  
TGCGCCCTGGGCGCCTCCAGCGGAGTGCCAGTACCCGGCTACTTCCGCCTGTGCGAGGCC  
TACGGCACGGCCAAGGGCTTCGGCAGTGGCGGCGGGCGGCACGCAGCAGCTCGCTAGTCC  
CTTTCCTGCGCAGCCCCCGGGGCGCGGTTTCGACCCGCCGCCCGCACTGGCCTCTGGCTC  
GACCGAGGCAGCCGGGAAGGAGCGAGTCCTAGACTCCACGCCACCACCCACTCTGGTTT  
GCACCGGTGGCGGCGGCTCGCAGGGCGACGAGGAGGCACACGCGTCATCCTCGGCGGCT  
GAGGAGCTGTCTCCAGCCCCTTCAGAAAACAGTAAAGCTTCGCCGGAGAAGGACTCCCTG  
GGCAGTTCCAAAGGCGAAAATGCAGCCAACTGGCTCACAGCAAAGAGCGGCCGGAAAGAA  
ACGCTGCCCTTACACGAAGCACCAGACGCTGGAGCTGGAGAAGGAGTTTCTATTCAACAT  
GTACCTTACTCGAGAGCGGCGCCTAGAGATCAGCCGTAGCGTCCACCTCACGGACAGACA  
AGTGAAAATCTGGTTTCAGAATCGCAGGATGAACTGAAGAAAATGAACCGAGAAAACC  
GAATCCGGGAGCTCACAGCCAACTTTAATTTTCTCTGA

### **Hoxa10ΔC1**

GGATCCTGCTCGGAGAGCCCTGCCGCGAACTCCTTTTTGGTCGACTCGCTCATCAGCTCAG  
GCAGAGGCGAGGCTGGTGTGGTGGCGGTAGCGCGGGGGGCGGTGGAGGTGGCTACTAC  
GCCACGGTGGGGTCTACCTGCCGCTGCCAGCGACCTGCCCAGCGCCGTCGCCCGGA  
GGCGGTGGCGGTGGTGGCAGCGGGGGCCTGGGTCTGGGACGCATGGCTACGCGCCCCG  
CCCCCTAGACCTGTGGCTGGACGCGCCCCGCTCCTGCCGGATGGAGCCGCCGACGGGCC  
GCCGCCACCGCAGCCACAACCCAGCAGCAGCAGCAGCAGCAGCCGCCGCCGCCGCCGCCG  
AGCCACCTCAACCCAGCCACAGGCCACTTCGTGTTCTTTTGCGCAGAACATCAAAGAAG  
AGAGCTCCTACTGCCTCTACGATGCTGCGGACAAATGCCCCAAGGGCTCGGCCGCCGCTG  
ATCTGGCCCCCTTTCCCGCGGGGGCCCGCCGCCGACGGCTGCGCCCTGGGCGCCTCCAGCG  
GAGTGCCAGTACCCGGCTACTTCCGCCTGTGCGAGGCCTACGGCACGGCCAAGGGCTTCG  
GCAGTGGCGGCGGCGGCACGCAGCAGCTCGCTAGTCCCTTTCCTGCGCAGCCCCCGGGGC  
GCGGTTTCGACCCGCCGCCGCCGCACTGGCCTCTGGCTCGACCGAGGCAGCCGGGAAGGAG  
CGAGTCCTAGACTCCACGCCACCACCCACTCTGGTTTGCACCGGTGGCGGCGGCTCGCAG  
GGCGACGAGGAGGCACACGCGTCATCCTCGGCGGCTGAGGAGCTGTCTCCAGCCCCTTCA  
GAAAACAGTAAAGCTTCGCCGGAGAAGGACTCCCTGGGCAGTTCCAAAGGCGAAAATGC  
AGCCAACTGGCTCACAGCAAAGAGCGGCCGGAAGAAACGCTGCCCTTACACGAAGCACC  
AGACGCTGGAGCTGGAGAAGGAGTTTCTATTCAACATGTACCTTACTCGAGAGCGGCGCC  
TAGAGATCAGCCGTAGCGTCCACCTCACGGACAGACAAGTGAAAATCTGGTTTCAGAATC

GCAGGATGAACTGAAGAAAATGAACCGAGAAAACCGAATCCGGGAGCTCACAGCCAA  
CTTTAATTTTTCTGA

### Hoxb9insC1

The C1 motif is shown in bold and underlined.

GGATCCATTTCTGGGACGCTTAGCAGCTATTATGTCGACTCGATCATAAGTCCCT**TACGGG**  
**CTGCAAAGCTGCGGGCTCTTCCCCGCGCTGGGCAGCAAGCGTAAT**GAAGCGCCGTCG  
CCCCACGAGAGCGAGGACGCGCCTCCAGCCAAGTTTCCTTCTGGCCAGTACGCGAGCTCG  
CGGCAGCCGGGCCACGCGGAGCACCTGGAGTTCCCCTCGTGCAGCTTCCAGCCCAAAGCG  
CCGGTGTTCGGCGCCTCCTGGGCGCCGCTGAGCCCGCACGCGTCCGGGAGCCTGCCGTCC  
GTCTACCACCCTTACATCCAGCCCCAGGGCGTCCCGCCGGCCGAGAGCAGGTACCTCCGC  
ACCTGGCTGGAGCCGGCGCCGCGCGGCGAAGCGGCCCGGGCAGGGCCAGGCGGCGGT  
GAAGGCGGAGCCGCTGCTGGGCGCGCCTGGGGAGCTGCTCAAACAGGGCACGCCCCGAGT  
ACAGTTTGAAACTTCGGCGGGCAGGGAGGCCGTGCTGTCTAATCAAAGACCCGGCTAC  
GGGGACAATAAAATTTGCGAAGGAAGCGAGGACAAAGAGAGGCCGGATCAAACCAACC  
CCTCCGCCAACTGGCTGCACGCTCGCTCTTCCCGGAAAAAGCGCTGTCCCTACACCAAAT  
ACCAGACGCTGGAGCTAGAGAAGGAGTTTCTGTTCAATATGTACCTCACCAGGGACCGTA  
GGCACGAAGTGGCCAGACTCCTCAATCTGAGTGAGAGACAAGTCAAAATCTGGTTTCAG  
AACCGGCGGATGAAAATGAAGAAAATGAATAAGGAGCAGGGCAAAGAGTA

### RNA probes used in this work:

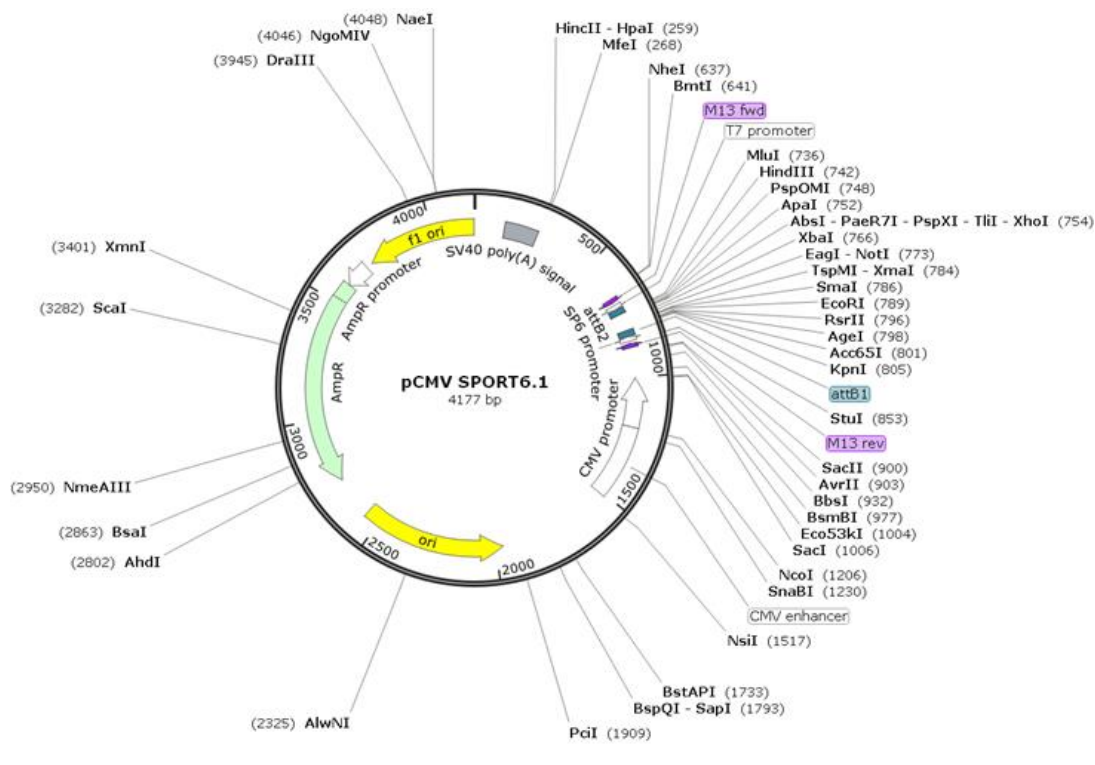
Name	Origin	Linearized with	Polymerase	Plasmid
<i>Tbx18</i>	Offered by Andreas Kispert	HindIII	T7	pBluescript II KS
<i>Uncx 4.1</i>	Offered by Bernhard Herrmann	SalI	T7	pSV-Sport1
<i>Paraxis</i>	cDNA	BamHI	T7	pBluescript II KS
<i>Myf5</i>	Offered by Jaime J. Carvajal	XbaI	T7	pBluescript II KS
<i>IFT144</i>	cDNA	SpeI	T3	pBluescript II KS

**Primer sequences used in this work:**

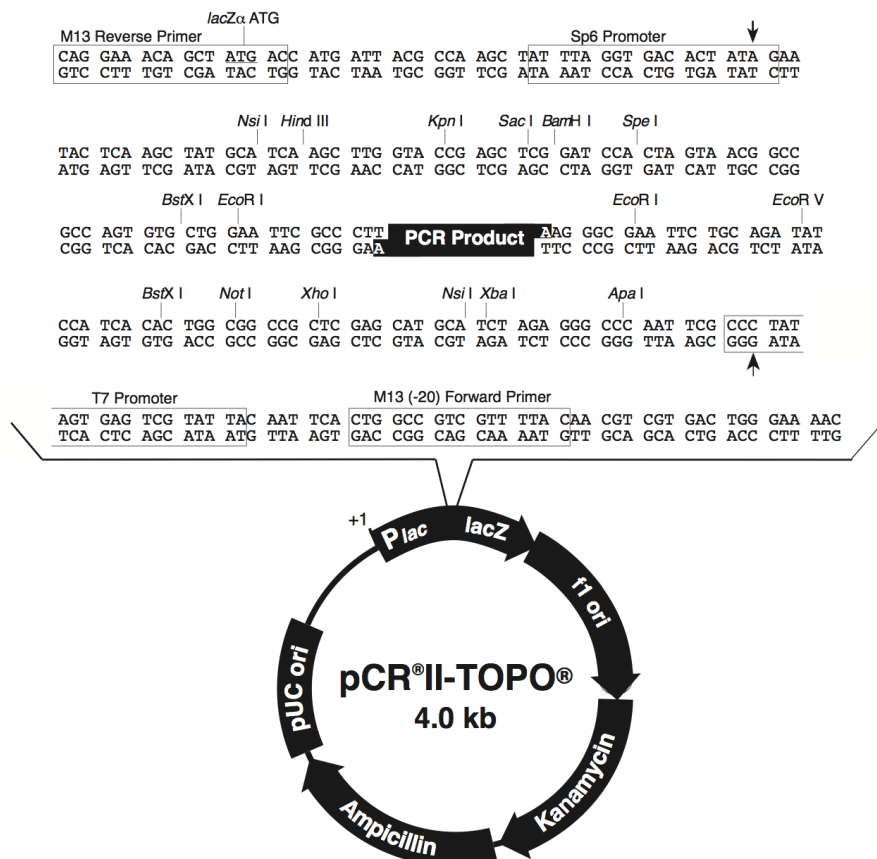
	Name	Sequence	Restriction Site	Orientation
<b>Grg3 KO Genotyping</b>	3322	TGTGCCCACTCTCTCCCCTATG	-	Forward
	3325	AAGGCCGCCCTCTCGCCAGC	-	Reverse
<b>WD40 Domain Cloning</b>	3348	CGGGATCCGATGGGCAGATGCAACCTGTGC	BamHI	Forward
	3349	CGGCGGCCGCGGGCCACGTGGGTGTCTCTC	NotI	Reverse
	3438	GATCTCCAAAGAAGAAGCGGAAGGTCGGTG	-	Forward
<b>IFT144 Cloning</b>	3490	GCGGATCCATGAAGCGTGTTTTCTCCCTGCT AGA	BamHI	Forward
	3491	GCCTCGAGATCACTAGTTAGGGCGTGGCATA A	SpeI/XhoI	Reverse
	3492	GCCCTAACTAGTGATTTCCCTCATCTACGG	SpeI	Forward
	3493	GCCTCGAGATTATCACCGGTGATGCCCTTCT CG	AgeI/XhoI	Reverse
	3503	GCACTAGTTCATTACGAGAAGGGCATCACC GGTGATAATAAGGA	AgeI/SpeI	Forward
	3512	GCCTCGAGGCGTACATACTGAAAAGAACAT CGTGTGCATTCCG	BsmI/XhoI	Reverse
	3516	GCGCTCTAGAGCAGGAACTATCGGAATGC ACACGATGTTCTT	BsmI/XbaI	Forward
	3517	GCTCTAGAGCCTCGAGGAGAGGTGCCATCT ACTCGGGAGC	XhoI/XbaI	Reverse
<b>IFT144 Probe</b>	3496 (with 3492)	CGCTGTCCCAGTGCTGCAGGTCC	PstI	Reverse
<b>DllB9insC1 Genotyping</b>	2019	AACTGGCTGCACGCTCGCTCTTCC	-	Forward
	2228	GGGAAGAGCTAGGGAGGACTG	-	Reverse
<b>DllHoxa10AC1 DllHoxa10AC1p1 DllHoxa10AC1p2 Genotyping</b>	1469	AATTCTGACCACCATGGATTATAAAGATGAC GATGACAAGG	-	Forward
	754	GTCCGTGAGGTGGACGCTACG	-	Reverse
	753	AGCGAGTCCTAGACTCCACGC	-	Forward

# **Appendix IV: Vector Maps**

pCMV Sport6.1



pCRII-TOPO



## pBluescript II KS (+/-)

