UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA ANIMAL

Analysis of Hox10 specific peptide motifs in their patterning functions of the axial skeleton

Isabel Misteli Guerreiro

Mestrado em Biologia Evolutiva e do Desenvolvimento
2009

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA ANIMAL

Analysis of Hox10 specific peptide motifs in their patterning functions of the axial skeleton

Isabel Misteli Guerreiro

Dissertação orientada por:

Professor Doutor Eduardo José de Frias Gonçalves Crespo
Doutor Moisés Mallo Perez

Mestrado em Biologia Evolutiva e do Desenvolvimento
2009

Acknowledgments

São muitas as pessoas que tiveram uma contribuição essencial para a realização desta tese.

Ao Moisés, que me deixou encontrar o meu lugar no seu laboratório e perceber o valor da curiosidade, da independência e da perseverança para ter sucesso no mundo da Ciência. Por continuar a acreditar em mim e dar valor ao meu trabalho. Porque o que aprendi com ele foi muito mais do que apenas técnicas laboratoriais.

Ao Professor Eduardo Crespo, por ter aceitado ser meu orientador interno.

Ao Instituto Gulbenkian de Ciência, uma instituição com excelentes condições e espírito científico.

A todas as pessoas com quem trabalhei lado a lado todos os dias e com quem partilhei frustrações, risos, parvoíces, cervejas, conversas com maior ou menor valor científico.
À Vanessa, a "minha colega de mesa", que todos os dias me dava um exemplo de força e boa disposição. Agradeço-lhe pelo apoio, pela compreensão, pelos desabafos. Desejo-lhe muita sorte para ela e a sua linda família.
À Jen, por me ter ensinado tanto, pelas conversas longas, pela amizade e pela ajuda. Thank you for being patient and explaining everything so well. You helped me more than you know.
Ao Arnon pelo bom humor brasileiro, pelo encorajamento, pelos cafés, pela companhia até às tantas. Pelas brincadeiras e pelos momentos divertidos "Cara!". Até mesmo pelas piadas de portugueses e loiras.
À Mafalda que, apesar de não ter ficado muito tempo, conseguia pôr um sorriso na cara de toda a gente assim que chegava. Por estar cheia de vida e por ser um "bicho social".
À Tânia, pelo apoio, por ter começado o trabalho e ter garantido que eu entrava com o pé direito na minha tese. Por se disponibilizar sempre para ajudar. Pela companhia e conselhos no "meeting" em Carmona.
À Ana, por tratar dos ratinhos e pelo excelente trabalho que desenvolve na unidade de transgénicos. Por ter sido um pouco responsável por cada um dos fenótipos que obtive.
Ainda que não tenha estado muito presente, à Filipa pelo seu jeito tão bem-disposto e particular. Pela sua entrega às artes dramáticas.

Ao grupo da organogénese, com o qual dividíamos a ala … e reagentes. À Raquel, pelas noitadas e pelas gargalhadas. Pelos "Domingos" e os lanches. Porque o laboratório ficava sempre mais triste quando ela voltava a Badajoz. À Rita por estar sempre lá tanto dentro como fora do laboratório, pela companhia e pelos desabafos. Pelo apoio incondicional e pela amizade. À Joana pelas "manias", pelas piadas e por tentar pegar fogo ao laboratório.

Ao "grupo de estudo", por ter tornado a escrita da tese divertida. À Rita e à Célia pela comida de plástico, gargalhadas e entreajuda. Pela amizade e muitos, muitos cafés.

Aos amigos e à família sem os quais teria sido difícil manter a minha sanidade mental.

Abstract

Hox genes play a fundamental role in anterior-posterior patterning and are remarkably conserved throughout evolution (Slack et al., 1993). Their products are transcription factors that regulate a specific set of genes with essential functions in development. Although different Hox genes show a notable functional specificity *in vivo*, they demonstrate a surprisingly low DNA-binding specificity *in vitro*. Sequence analysis can provide a way to understand how Hox genes achieve their biological specificity (Prince, 2002).

Genetic experiments revealed that Hox genes are involved in global patterning processes in the axial skeleton to produce the axial formulae. Hox group 10 genes, in particular, have been shown to repress thoracic rib formation, since their overexpression in the presomitic mesoderm causes a ribless phenotype and their global inactivation resulted in extra ribs (Wellik *et al.*, 2003, Carapuço *et al.*, 2005).

Two peptide domains were identified in Hox10 proteins which are conserved among all the Hox 10 members and are absent from all other Hox proteins. One of these is an octapeptide located just N-terminal to the homeodomain. The purpose of this work is to understand the role of this octapeptide in Hox10 protein function. This is being approached by the genesis and functional analysis of transgenic mice expressing mutant Hoxa10 proteins that contain specific deletions or amino acid changes in this domain. In previous transgenic assays, the overexpression of *Hoxb9* gene in the presomitic mesoderm did not produce an abnormal axial skeleton phenotype. For this reason, this gene was used to generate chimeric contructs with the *Hoxa10* gene.

The results obtained show that the removal of the octapeptide is sufficient to block the rib-repressing activity of Hoxa10 when expressed in the presomitic mesoderm. In addition, introduction of this peptide motif, as well as the whole Hoxa10 sequence N-terminal to it, into the Hoxb9 protein produced a partial ribless phenotype. These results indicate that the octapeptide is necessary for the rib-repressing activity of Hoxa10 but it does not seem to be sufficient for this function, at least individually.

Keywords: Hox genes, axial patterning, functional specificity, protein sequence

Resumo

Em mamíferos, existem 39 genes Hox responsáveis por especificar a polaridade antero-posterior (AP). Estes genes são homólogos dos genes selectores homeóticos que especificam a identidade segmentar em *Drosophila melanogaster.* Nos vertebrados, tal como em Drosophila estes genes são expressos pela mesma ordem pela qual estão distribuídos nos cromossomas. Os genes Hox dos mamíferos demonstram não só esta colinearidade a nível espacial como também a nível temporal, sendo que os genes que se encontram mais a montante no cromossoma são também os que começam a ser expressos em primeiro lugar (Duboule, 1998, Duboule *et al.*, 1989). Para além da sua expressão ser regulada pelos seus próprios produtos, os genes Hox são também directamente regulados pelos genes Cdx. As proteínas Fgf, Wnt e o ácido retinóico regulam também, de forma directa ou indirecta, a expressão de genes Hox (Deschamps *et al.*, 2005).

As proteínas Hox são responsáveis pela regulação de vários genes envolvidos em diversas funções essenciais do desenvolvimento animal, incluindo a adesão, ritmos de divisão, morte e movimento celular (Favier *et al.*, 1997).

Os factores de transcrição codificados por estes genes têm um elemento altamente conservado de ligação ao DNA ao qual se dá o nome de homeodomínio (HD) (Gehring et al., 1990). O HD é constituído por três hélices-α e uma extensão N-terminal adjacente à primeira hélice. A terceira hélice-α reconhece uma sequência de DNA composta por seis pares de base com um núcleo conservado de quatro pares de base. No entanto, esta sequência que o HD reconhece nos seus genes-alvo surge com relativa frequência no genoma, não parecendo, por isso ter uma especificidade de ligação ao DNA muito elevada (Ekker et al., 1991). Assim, apesar destes genes demonstrarem uma elevada especificidade funcional *in vivo*, a sua especificidade de ligação ao DNA *in vitro* demonstrou ser muito baixa. Uma vez que o homeodomínio mostra uma elevada conservação entre os vários grupos de genes Hox, é pouco provável que este seja suficiente ou mesmo necessário para conferir as funções específicas de cada grupo parálogo. A análise das sequências e estrutura das proteínas Hox aponta cada vez mais para a possibilidade da especificidade biológica ser essencialmente conferida por resíduos que não contactam com o DNA (Merabet *et al.*, 2009, Merabet *et al.*, 2003, Sharkey *et al.*, 1997). Algumas proteínas Hox mostram poucos aminoácidos conservados dentro do HD em todo o grupo parálogo, indicando que a sua especificidade funcional poderá residir essencialmente em sequências exteriores ao HD (Sharkey et al., 1997).

Os modelos sugeridos indicam para a possibilidade da ligação a outras proteínas (co-factores) ser responsável por despoletar a regulação dos genes a jusante ou por aumentar a especificidade e afinidade de ligação dos factores de transcrição aos elementos cis-regulatórios (Biggin *et al.*, 1997, Li *et al.*, 1999).

Experiências genéticas já mostraram que os genes Hox são importantes em vários aspectos do desenvolvimento animal, incluindo na modulação do esqueleto axial (Wellik, 2007). Foi, inclusivamente, demonstrada uma correlação entre a expressão anterior de genes Hox e os limites de regiões morfologicamente distintas do esqueleto axial (Burke et al., 1995). Este deriva de estruturas mesodérmicas que ladeiam o tubo neural denominadas sómitos (Dale *et al.*, 2000).

Em ratinho, os genes Hox do grupo 10 mostraram impedir a formação de costelas torácicas. Ao impedir a expressão deste gene formam-se costelas torácicas na zona lombo-sacral (Wellik *et*

al., 2003). Por outro lado, a sua sobre-expressão na mesoderme pré-somítica gera esqueletos sem costelas (Carapuço et al., 2005).

Este estudo tem como objectivo entender que aminoácidos são necessários para conferir a especificidade funcional aos genes *Hox10* no controlo da formação do esqueleto axial.

Foram identificados dois domínios peptídicos nas proteínas Hox do grupo 10 que apresentam uma conservação perfeita restringida apenas a este grupo de proteínas. Um destes é um octapéptido que se localiza a N-terminal do homeodomínio. Para perceber a contribuição deste octapéptido na função das proteínas Hox do grupo parálogo 10, foram geradas construções diferentes em que o cDNA de *Hoxa10* foi alterado utilizando técnicas básicas de clonagem. As proteínas mutantes e quiméricas foram criadas por uma técnica de mutagénese por PCR. Depois de clonados, os cDNAs foram microinjectados de forma a gerar ratinhos transgénicos que foram recolhidos no dia embrionário 18.5. O seu esqueleto foi então corado e analisado em detalhe. No total, foram geradas quatro construções mutantes. De forma a perceber se este domínio é necessário à função das proteínas Hox10, o octapéptido foi removido. Dos esqueletos obtidos de embriões transgénicos a maioria tinha um fenótipo normal. Apenas um apresentou defeitos de fraca intensidade no esqueleto axial, nomeadamente a nível do esterno. Este motivo peptídico parece, por isso, ser necessário à função das proteínas Hox10. Dois aminoácidos pertencentes a este domínio peptídico foram também modificados para uma análise mais detalhada do octapéptido. Os esqueletos resultantes da sobre-expressão deste cDNA modificado não mostraram qualquer tipo de deficiências ao nível do esqueleto axial. Este resultado poderá indicar que a fosforilação destes dois aminoácidos confere a actividade repressora da formação de costelas torácicas. Para confirmar que esta proteína estava a ser produzida, a construção mutada foi transfectada para duas linhas celulares distintas. Em ambas as situações a produção de proteína foi confirmada, apesar de uma das linhas celulares ter gerado proteínas mais pequena do que o esperado.

O gene *Hoxb9* já tinha mostrado não ter fenótipo a nível do esqueleto axial quando sobre-expresso na mesoderme pré-somítica (Laboratório de M. Mallo, dados não publicados). Para determinar se o octapéptido é suficiente para conferir a funcionalidade das proteínas Hox10, este domínio peptídico foi inserido na proteína Hoxb9. Ao sobre-expressar esta construção obtiveram-se transgénicos com defeitos no esqueleto axial, incluindo deficiências na ligação das costelas ao esterno e alteração da morfologia das vértebras. No embrião com o fenótipo mais grave foi também observada a ausência de costelas em três vértebras torácicas. Estes resultados indicam que o octapéptido tem um papel relevante na especificidade funcional das proteínas Hox10 não conseguindo, no entanto, reprimir completamente a formação de costelas torácicas.

De forma a explorar outros aminoácidos que pudessem, juntamente com o octapéptido, ser suficientes para conferir a especificidade funcional às proteínas Hoxa10, foi gerada outra proteína quimérica. Esta proteína inclui toda a parte da sequência de Hoxa10 a montante do octapéptido e, a jusante deste, a sequência é exclusivamente de Hoxb9. Os fenótipos transgénicos obtidos não foram significativamente diferentes dos gerados pela construção que continha apenas o octapéptido. Estes resultados parecem indicar que os aminoácidos de Hoxa10 adicionados à construção não têm impacto detectável na função normal das proteínas Hox9.

Este estudo sugere que as sequências fora do HD têm um papel extremamente importante na função dos genes Hox. Observados na sua globalidade, estes resultados parecem indicar que o

octapéptido é necessário para a função repressora de costelas torácicas do Hoxa10. No entanto, este domínio peptídico não parece ser suficiente para levar a cabo esta função, pelo menos não por si só. Outros aminoácidos, dentro ou fora do HD, parecem ser necessários para este factor de transcrição regular correctamente os seus genes-alvo. O uso de proteínas mutantes ou quiméricas é um importante passo inicial para compreender os mecanismos pelos quais os genes Hox exercem a sua especificidade funcional e, consequentemente, estabelecem diferenças ao longo do eixo AP.

Palavras-chave: Genes Hox, padronização axial, especificidade funcional, sequências peptídicas

Table of contents

I. **Introduction**

I.1 Hox genes

In 1984, William Bateson coined the term "homeosis" as a type of variation in which certain body parts assumed a morphological identity similar to that of another region (Bateson, 1984). Later, mutations in a group of genes initially identified in the fruit fly *Drosophila melanogaster*, was found to be responsible for homeotic transformations of specific segments along the anterior-posterior (AP) axis. These genes were thus designated homeotic genes (Lewis, 1978, Gehring, 1987).

In *Drosophila melanogaster*, there are eight clustered Homeotic genes that control segmental identity along the AP axis: *labial* (*lab*), *proboscipedia (pb), bicoid (bcd)*, *Deformed (Dfd)*, *Sex combs reduced* (*Scr)*, *Antennapedia (Antp), Ultrabithorax (Ubx), Abdominal-A (Abd-A), Abdominal-B (Abd-B)*. These are grouped in two complexes located in the third chromosome: the Bithorax (BX-C) and the Antennapedia (ANT-C) complexes (Fig 1a)(Lewis, 1978).

The homeotic genes are characterized by the presence of a conserved 183 base pair DNA sequence, the homeobox, which encodes the homeodomain (HD). This 61 aminoacid motif has DNA binding ability and is essential for Hox proteins to function as transcription factors, regulating the expression of genes with important roles in development (Gehring et al., 1990).

Homeobox-containing genes were later identified in many metazoan genomes by low stringency screening, including mice (Hart et al., 1985), frogs (Carrasco et al., 1984) and humans (Boncinelli et al., 1985). It now seems likely that a linked cluster of Hox genes is a character shared by all metazoan and may be fundamental to control axial patterning in animals (Slack et al., 1993).

In mammals there are 39 Hox genes that have been shown to have sequence and structural similarities to the homeotic selector genes found in *Drosophila*. Contrary to arthropod homeotic genes, mammalian *Hox* genes are organized in four clusters (A, B, C, and D) located in four different chromosomes. These appear to have originated from an ancestral single cluster through whole-genome duplication events. Within each cluster, Hox genes can be subdivided into 13 sets of genes called paralogous groups based on their homology to the homeotic *Drosophila* genes and to each other. However, not all paralogs are represented in each cluster probably as a result of secondary gene losses (Fig 1b) (Prince, 2002).

All animals exhibiting AP axial polarity have their Hox genes organized in the chromosome in such a way that reflects the position in the developing body axis where their transcription is activated. In mammals, *Hox* genes closer to the 3' extremity are activated first and in more anterior domains, whereas 5'-located genes are transcribed later and in more posterior areas. This phenomenon was termed spatial and temporal collinearity (Duboule *et al.*, 1989, Duboule, 1998). However, the temporal aspect of the colinearity is only shared by vertebrate and short germ insect species (Duboule, 1998). Another interesting feature that characterizes Hox genes is that, both in *Drosophila* and in mice, they all have the same transcriptional orientation.

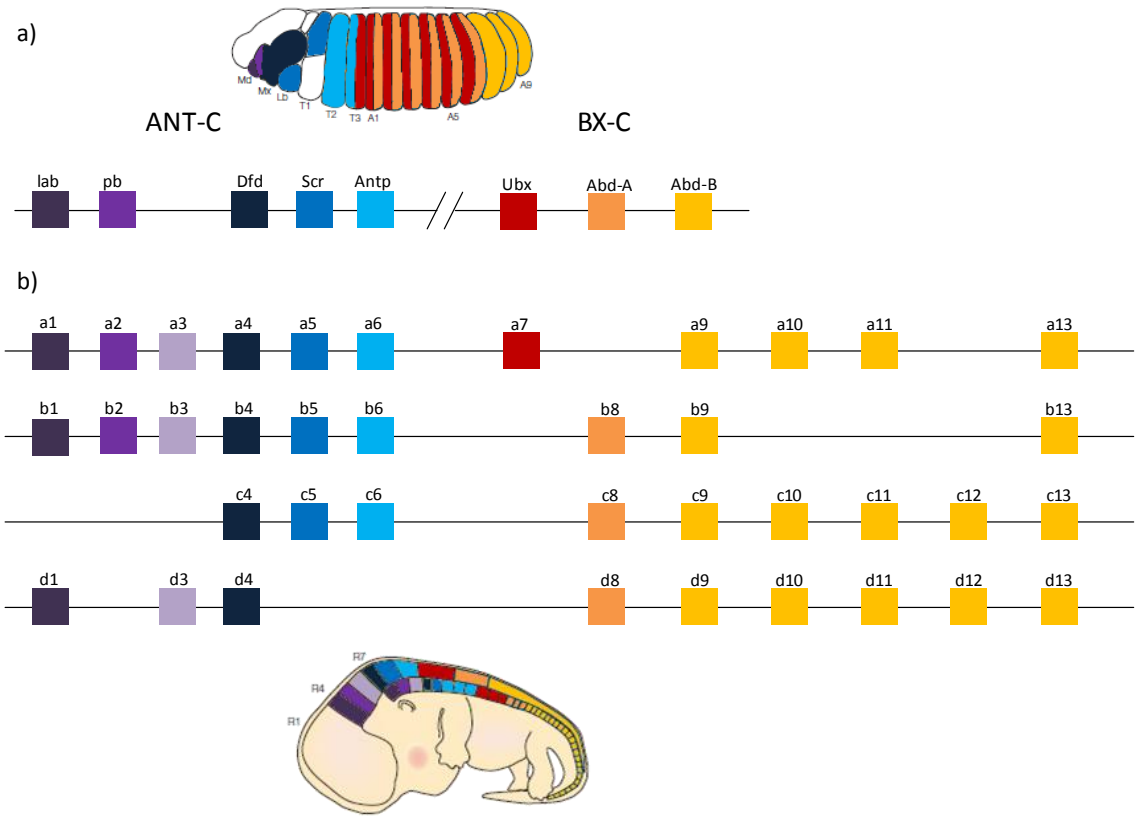


Figure 1 Hox gene expression and genomic organization in Drosophila and mouse.
a) Schematic representation of a Drosophila embryo (top) with colors showing the approximate domains of expression. Bellow, the gene distribution along the chromosome and the Antennapedia and Bithorax complexes are represented. b) The paralogous goups within the four clusters are color-coded according to their assumed phylogenetic relationship with the Drosophila Hox genes. At the bottom, a mouse embryo is represented and each color illustrates the anterior-most expression domain of each subfamily of Hox genes. Adapted from Pearson et al., 2005.

I.1.1 Hox gene expression

Mouse genes, located in the 3' extremity of the clusters, start to be expressed in the mesoderm and more weakly in the primitive ectoderm of the embryo's posterior primitive streak (Gaunt, 1988, Gaunt *et al.*, 1994, Deschamps *et al.*, 1999). Transcription initiation of more 5' genes occurs in the same region, at progressively later stages where new epiblast cells have been brought by gastrulation movements. In general, after this initial expression, Hox transcripts spread in a posterior to anterior movement until they reach their most rostral position (Tam *et al.*, 1987). Gene expression levels continue increasing until embryonic day (E) 12,5 (Kessel *et al.*, 1991). By this time, collinear sharp anterior boundaries have been established, while transcript levels decline gradually as they reach the posterior end of the embryo. Consequently, there is an increase in overlapping regions and diversity of Hox gene expression in more posterior areas of the body (Favier *et al.*, 1997). However, these genes do not appear to act together to determine the patterning information. In fact, when a posterior gene is ectopically expressed in a more anterior domain, it also undergoes a homeotic transformation in which posterior-like structures are formed. Interestingly, the opposite does

not occur. This phenomenon, by which more posterior genes impose their function over that of Hox genes expressed in more anterior regions, is termed posterior prevalence (Duboule *et al.*, 1994, Kmita *et al.*, 2003).
Hox gene expression is found in several germ layer derivatives and has crucial functions in various developing systems such as the limbs, the developing hindbrain, the pharyngeal arches, the developing genito-urinary tract and the axial skeleton (Favier *et al.*, 1997). The role of Hox genes in the axial skeleton in particular will be approached in more detail for its relevance to the present study.

I.1.2 Hox gene regulation

The correct timing at which Hox genes are initially activated is crucial for the establishment of accurate expression domains. In this initial phase, Hox genes appear to be regulated by Wnt and Fgf proteins which have a role in the formation and morphogenetic movements through the primitive streak. In later stages, as Hox gene expression spreads further toward the anterior end, paraxial mesoderm cells are exposed to patterning signals. Either directly or indirectly, Fgf and Wnt proteins have all been shown to regulate Hox gene expression. Likewise, the presence of different levels of retinoic acid along the PSM seems to be essential for Hox regulation and successful axial patterning (Deschamps *et al.*, 2005).

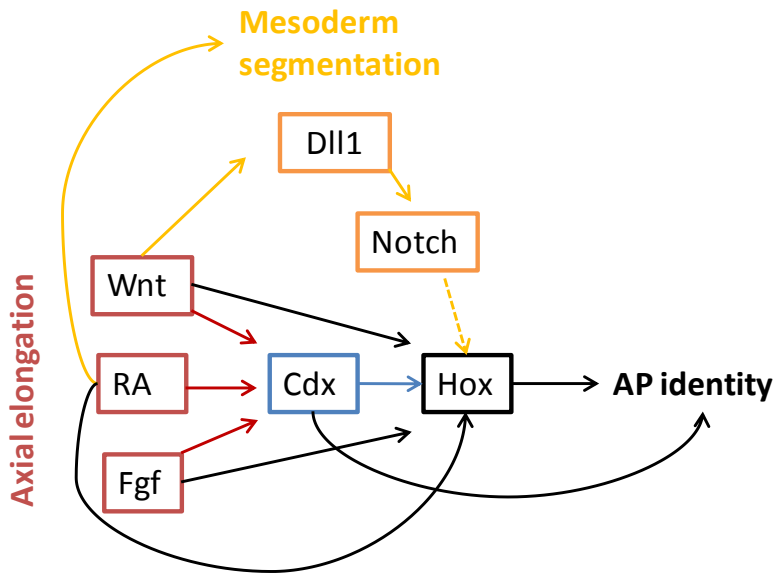


Figure 2 Hox gene regulation network.
Simple schematic representation of interactions between genes involved in axial elongation, somitogenesis and AP patterning. Adapted from Deschamps and Van Nes, 2005.

Hox gene expression is also affected by genes involved in the segmentation program. For instance, the loss-of-function of a Notch ligand (Delta-like 1) results in anterior homeotic transformations of the vertebrae as well as a posterior shift of Hox expression domains, providing evidence for the involvement of the Notch pathway in Hox gene regulation (Cordes et al., 2004).
The Cdx genes, a Hox-related family, have been shown to directly regulate Hox genes in the mesoderm and neuroectoderm in a dose-dependent manner. A summary of the different genes involved in Hox gene regulation and establishment of AP identity is illustrated in Fig 2.

Furthermore, Hox genes may be autoregulated by their own products or controlled by other Hox proteins, although the mechanism of this complex interaction remains largely undefined.
In addition to the regulation provided by these molecular signals, Polycomb (PcG) and trithorax (trxG) group proteins play an important role in maintaining Hox gene expression spatially restricted. These protein groups act by altering the transcriptional states of Hox genes through chromatin structure modifications (Deschamps *et al.*, 2005). The combined action of PcG proteins that maintain repression of Hox gene expression and trxG proteins that sustain it, ensures that Hox genes have their expression restricted to the correct areas (Mahmoudi *et al.*, 2001).

I.1.3 Hox target genes

Hox proteins, as monomers, heterodimers or part of larger complexes are responsible for the regulation of a large pool of genes. Apart from their ability to regulate themselves, as well as some of their known co-factors, Hox transcription factors also act on genes that mediate adhesion, cell division rates, cell death and cell movement. However, many target genes still remain to be discovered. Either directly or indirectly, Hox proteins clearly have a pivotal role in regulating most of the genes involved in animal development (Pearson et al., 2005).

I.2 Hox functional specificity

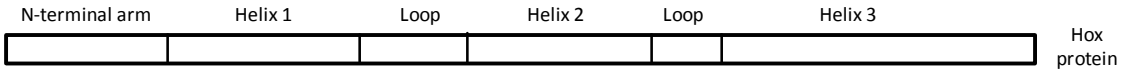
I.2.1 The paradox

The homeodomain structure has been determined both by nuclear magnetic resonance spectroscopy (Otting et al., 1990) and X-ray crystallography (Kissinger et al., 1990). It consists of three α-helices and a flexible N-terminal arm adjacent to the first helix (Fig 3a). The third helix, also called recognition helix, contacts the major groove of DNA and recognizes a six base-pair DNA sequence that contains a four base pair recognition core (Fig 3b)(Ekker et al., 1991). Shen *et al.* showed that, in the presence of Pbx, more anterior Hox proteins preferentially bind to a TGAT core sequence while Hox proteins 6-10 essentially recognize a TTAT core. Hox proteins 3-8 can also bind to a TAAT core sequence (Shen et al., 1997). However, this TNAT sequence is very common in the genome, which raises some very interesting questions. Homeodomain-containing proteins show a remarkable functional specificity *in vivo* and are obviously capable of activating or repressing the correct set of genes in the right place at the right time. So how can these transcription factors bind to the correct target sequences? How do they know which genes to activate or repress? And how come different homeoproteins act differently even though they have the ability to bind to very similar sequences?
Two models have been proposed to explain this paradox. In the first model the Hox protein is only able to attach to DNA when bound to other proteins (co-factors) (Biggin *et al.*, 1997, Li *et al.*, 1999). In fact, different Hox/cofactor heterodimers have in several cases shown distinct DNA-binding specificities (Chan *et al.*, 1996).
The other model, on the other hand, assumes that Hox proteins are already bound to several DNA sites as monomers. Co-factors will differentially bind to the attached transcription factors,

changing them to an activated state in order to allow the transcription of the correct target genes (Biggin *et al.*, 1997, Li *et al.*, 1999).

These models are not mutually exclusive. In fact, both cases have been reported and experimentally verified. It is clear that Hox proteins act using a complex array of molecular strategies to regulate the correct set of genes, which will originate distinct structures according to their position along the body axis and developmental timing.

a)



b)

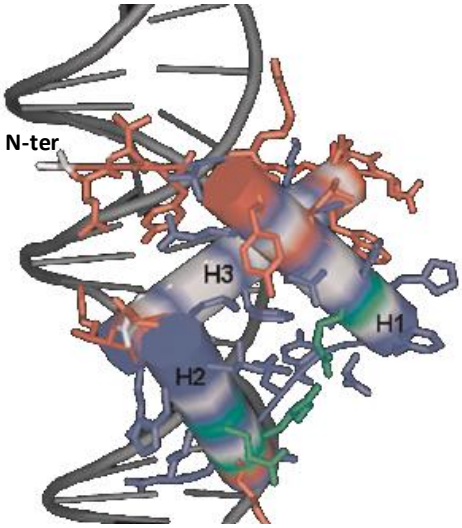


Figure 3 Hox protein structure.

a) Secondary structure representation of a Hox protein and (b) 3-dimensional structure of the Hox-DNA interaction adapted from Merabet et al., 2009. The N-terminal arm (N-ter), helix1 (H1), helix2 (H2) and helix3 (H3) are indicated.

I.2.2 Hox protein structure and sequence analysis

Sequence analysis of Hox proteins and identification of residues specific to a given paralog group have proven to be very important to clarify how Hox proteins achieve functional specificity *in vivo*.

When comparing Hox proteins with other homeodomain-containing proteins it was shown that most residues found in all Hox proteins are also conserved in most homeodomain classes. These common residues are either required for proper HD folding or essential for DNA binding. There are just four amino acids that can be classified as Hox generic signatures and are not found in other homeodomain protein classes. Only one of the amino acids that distinguish Hox proteins from other HD proteins has a DNA-binding function, whereas the other three are located in regions that could possibly interact with other proteins (Merabet et al., 2009).

UNIQUE PARALOG SIGNATURES

The mutation of a single paralogous group member usually yields a very mild phenotype, particularly in the axial skeleton. However, when more genes from the same paralogous group are mutated, a synergistic effect occurs and a much serious phenotype is observed. Therefore, the members of a given paralogous group seem to have a great deal of redundancy in their

function (Wellik *et al.*, 2003, Fromental-Ramain *et al.*, 1996). Thus, residues that are conserved in all members of a paralogous group but are not observed in others could clarify which amino acids are necessary for the functional specificity of Hox proteins (Sharkey et al., 1997).
The comparison of residues exclusively conserved in the HD of each paralog group revealed that most of the unique paralog signature residues are in positions favorable for protein-protein interactions (Merabet et al., 2009). On the other hand, the ones that contact DNA are mostly located in the N-terminal arm, which has been reported to play a major role in providing functional specificity to Hox proteins. It has been recently suggested that, while the conserved third helix of different paralogs recognizes similar binding sites, the N-terminal arm residues bind in a more specific manner by recognizing the structure and electrostatic potential of the minor groove (Fig 3b) (Joshi et al., 2007).
In order to determine the contribution of unique paralog residues to Hox protein function, chimeric Hox protein experiments have been performed, in which these specific residues were swapped by those of another Hox protein (Zeng *et al.*, 1993, Furukubo-Tokunaga *et al.*, 1993, Lin *et al.*, 1992, Chauvet *et al.*, 2000, Joshi *et al.*, 2007). Even though the HD has proven to have a critical role in DNA-binding specificity *in vivo*, some Hox proteins lack features that distinguish them from members of other paralog groups. For this reason, it is to be expected that peptide motifs outside the HD might have a crucial role in the establishment of characteristic Hox protein function. It was found that paralogs 12 and 13 have very few or no conserved aminoacids outside the HD. However, Hox proteins 1-8 all share a hexapeptide motif with a conserved YPWM core motif, located N-terminal to the HD. Paralog groups 9, 10 and 11 have conserved aminoacids immediately adjacent to the HD. Hox9 has characteristic residues N-terminal to the HD, while Hox11 has conserved residues C-terminal to the homeodomain. Hox10 signatures, on the other hand, are found both upstream and downstream the HD (Sharkey et al., 1997). More specifically, Hox10 proteins show a conserved NWLTAKSG octapeptide motif adjacent and N-terminal to the HD, as well as a RENRIRELT motif just C-terminal to the HD (Fig 12a).

I.2.3 Hox co-factors

Results obtained by sequence analysis strongly suggest that biological specificity is mostly achieved through protein-protein interactions, since many of the paralog group characteristic residues do not contact DNA (Chauvet *et al.*, 2000, Sharkey *et al.*, 1997). In fact, Hox proteins have already been reported to be unable to achieve functional specificity solely by DNA-protein interactions (Schier *et al.*, 1993). Most Hox co-factors identified so far are homeoproteins with Hox-independent functions that belong to the TALE as well as the POU family. The best studied and well known co-factors are part of the TALE family of homeodomain proteins, which is characterized by the presence of a three-amino acid extension in the loop between helices 1 and 2 of their HD (Mann *et al.*, 1998, Moens *et al.*, 2006).
The first Hox co-factor was identified in Drosophila and termed Extradenticle (Exd). The loss of function of this gene produced mutations in embryonic pattern without altering Hox gene expression. Later, a vertebrate homolog called Pbx1 was independently discovered as the cause of human preB cell acute lymphoblastic leukemia. Exd and Pbx proteins are included in the PBC subclass of TALE proteins which have highly characteristic and conserved peptide

domains N-terminal to the HD (Mann *et al.*, 1996, Chan *et al.*, 1996). Proteins from the PBC class have the ability to bind to the conserved hexapeptide observed in Hox proteins 1-8. Initially it was reported that Abd-B class proteins were unable to bind to PBC protein class members but this was later found to be due to the use of the wrong target DNA (Chang et al., 1995). Hoxa10 and Hoxb9, which lack a YPWM motif, have been shown to form a DNA binding complex with Pbx1, mediated by a motif that comprises a conserved tryptophan located just upstream to the homeodomain (Chang *et al.*, 1996, Shen *et al.*, 1997). However, members of the 11, 12, and 13 paralogs were found to be unable to bind DNA with Pbx (Shen et al., 1997).
PBC proteins cooperatively associate with Hox proteins via the three-amino acid loop in the homeodomain, apparently increasing Hox-binding selectivity (Burglin, 1997, Chan *et al.*, 1996). The heterodimer binds to DNA via a bipartite sequence and can either activate or repress its target genes (Moens *et al.*, 2006). However, not all Hox proteins seem to require Pbx for higher DNA-binding specificity (LaRonde-LeBlanc *et al.*, 2003).
The MEIS class is another well known class of TALE proteins that have the ability to bind to PBC proteins and may also bind directly to the Hox protein, (Burglin, 1997). In vertebrates, this class of homeoproteins includes Meis and Prep proteins and are involved in nuclear localization and stability of Pbx proteins (Moens *et al.*, 2006).
Extensive studies have been conducted on these co-factors in order to understand the exact mechanisms by which they are able to direct Hox DNA-binding (Mann *et al.*, 1996, Rieckhof *et al.*, 1997, Berthelsen *et al.*, 1998). However, many questions remain unanswered. For example, the fact that PBC proteins have the ability to bind to most Hox proteins does not clarify why they choose to bind to one particular Hox protein instead of any other expressed in the same region. Since all co-factors analyzed to date have Hox-independent function, it seems likely that Hox proteins bind to pre-existing complexes, providing them with spatial information to correctly control downstream genes (Mann *et al.*, 1998). Most components of the several possible complexes responsible for the regulation of the massive pool of Hox target genes have not been fully described. It is likely that more Hox-interacting proteins are yet to be identified. The Antennapedia YPWM peptide motif, for example, was reported to interact with a protein other than Pbx in order to cause the typical eye-to-wing transformation observed when Antennapedia is ectopically expressed (Prince et al., 2008). In addition Hox genes can bind to proteins such as histone acetyltransferases and histone deacetylases that modify chromatin structure, regulating downstream targets through epigenetic mechanisms (Shen *et al.*, 2001, Saleh *et al.*, 2000, Lu *et al.*, 2003). The presence of conserved residues characteristic of a given paralog group with no DNA-binding functions is a step forward toward the understanding of Hox functional specificity mechanisms. It is also possible that these conserved residues are necessary for structural reasons and have no direct influence in protein-protein interactions. In any case, new ways for Hox transcription factors to regulate developmental pathways could be uncovered.

I.3 Hox genes and axial skeleton

I.3.1 Somitogenesis

In the vertebrate embryo, the axial skeleton, skeletal muscle and dorsal dermis arise from transient mesodermal structures that lie on both sides of the neural tube called somites. Somites are sequentially produced from the anterior end of the unsegmented paraxial

mesoderm in a rostral to caudal direction. The formation of somites is precisely regulated and occurs at regular time intervals (Brent *et al.*, 2002, Dale *et al.*, 2000, Aoyama *et al.*, 1988). At the molecular level, the formation of somites includes two processes. One of them is characterized by the oscillating expression of a set of genes along the presomitic mesoderm. The chick gene Hairy1 was the first gene described to have this kind of "clock-like" behavior (Palmeirim et al., 1997). Since then, many genes have been found to have the same oscillatory behavior in a variety of vertebrate species (Dale *et al.*, 2000). These genes belong mostly to the Wnt and Notch signaling pathways. The second process required for somitogenesis is the segmentation signal, which seems to be regulated by the opposing gradients of Wnt/Fgf (posterior to anterior) and retinoic acid (anterior to posterior). As cells progress towards the anterior end of the PSM, they become less exposed to the Fgf signal. Eventually, cells reach an Fgf threshold where the segmentation program is activated. This "determination front" defines a change in gene regulation, ultimately resulting in the formation of somites (Mallo, 2007, Dubrulle *et al.*, 2004, Mallo *et al.*, 2008, Pourquie, 2003).

After its formation, molecules from surrounding tissues signal the somites to differentiate into compartments that will give rise to distinct cell lineages. The dorso-lateral part of the somite originates the dermomyotome, whereas the ventro-medial part de-epithelializes to form the mesenchymal sclerotome. The dermomyotome further differentiates into the dermotome, which gives rise to dermis, and the myotome, which gives rise to the axial musculature. The sclerotome undergoes a resegmentation process so that the posterior half of one somite and the anterior half of the next somite give rise to a single vertebral element. The cells in the ventro-medial sclerotome undergo an epithelial-to-mesenchymal transition and migrate to form the chondrocytes of the vertebrae and the proximal part of the ribs. Although most of the axial skeleton is derived from the somites, the sternum arises from the lateral plate mesoderm (Fig 4)(Brent *et al.*, 2002, Wellik, 2007, Gilbert, 2006, Monsoro-Burq, 2005, Aoyama *et al.*, 1988, Buckingham, 2001).

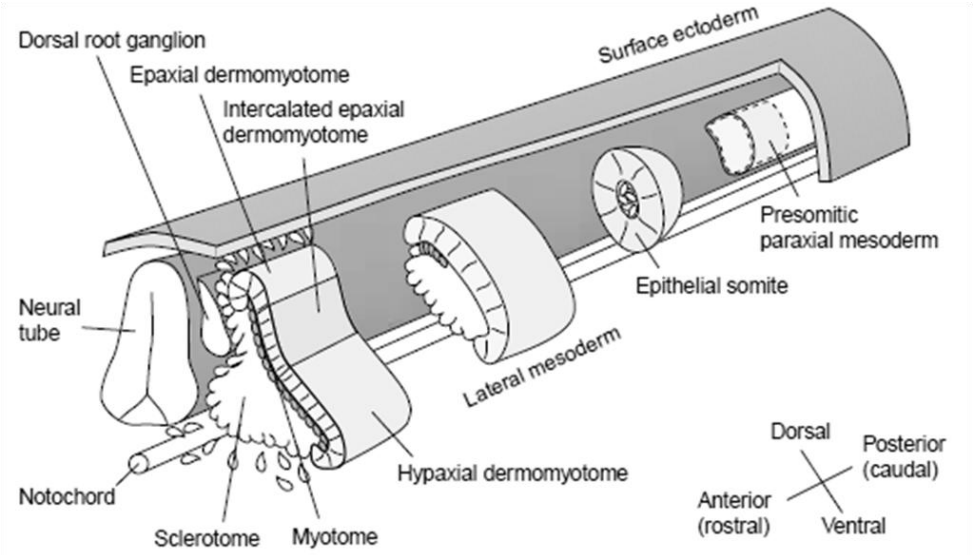


Figure 4 Schematic representation of vertebrate somitogenesis.
Somites mature following a posterior to anterior direction. From Buckingham,2001.

I.3.2 Hox and the axial skeleton

Although somites appear morphologically similar, they will differentiate into morphologically distinct vertebrae depending on the position they assume along the AP axis. The number of vertebrae in each morphological type (i.e. cervical, thoracic, lumbar, sacral and caudal), known as the axial formula, is largely determined by Hox gene expression. In mice, the axial formula is composed of seven cervical, thirteen thoracic (with ribs), six lumbar, four sacral and a variable number of caudal vertebrae (Burke et al., 1995). Thoracic vertebrae are characterized by the presence of ribs. Sacral vertebrae lack the fully formed ribs found in the thorax but they bear modified rib-like structures that fuse to form the sacrum (Fig 5).

The observation of anterior expression domains of many paralogous group *Hox* genes in both chickens and mice, demonstrated a correlation between these expression limits and the boundaries of morphologically distinct regions of the axial skeleton (Burke *et al.*, 1995, Burke, 2000). Accordingly, mutations of genes with an anterior expression limit close to a transition in vertebrae morphological type have confirmed this apparent correlation (Wellik, 2007).

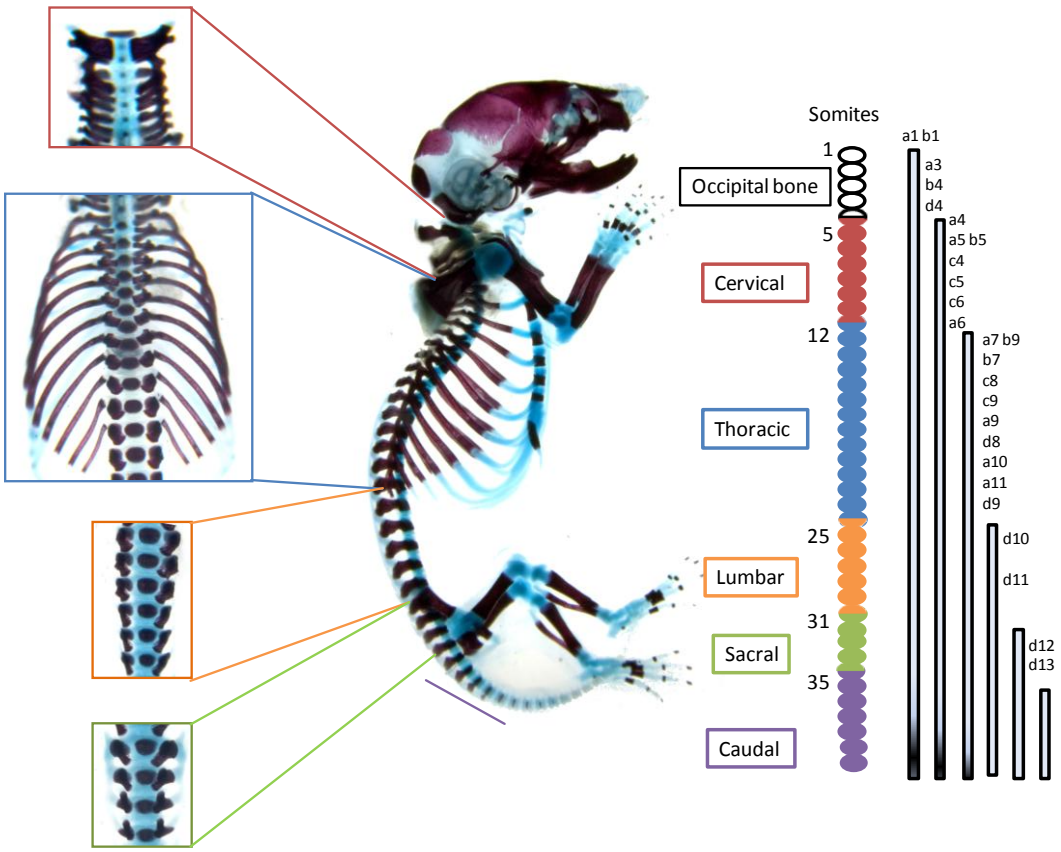


Figure 5 Representation of the different vertebrae types of a normal mouse embryo and somites that give rise to each of them.

At the right, Hox gene expression domains along the AP axis are illustrated in a simplified manner. The reducing gradient shows the general trend for decreased expression in more posterior domains. Adapted from Burke et al., 1995 and Favier and Dolle, 1997.

The fact that the axial skeleton is derived from two different primitive tissues, makes it hard to correctly interpret phenotypes at the thoracic rib cage level. While the vertebrae and the proximal part of the ribs originate from the somites, the sternum and the more distal part of the ribs are derived from the lateral plate mesoderm (McIntyre et al., 2007). It has been reported that the inactivation of entire paralogous groups resulted in anterior homeotic

transformations in the axial skeleton. Interestingly, mutations of different paralog groups varied in their affected areas in a collinear fashion. That is, more anterior genes are responsible for the patterning of more anterior axial structures, while mutants of more posterior genes influence more posterior structures. However, this is not observed in the abaxial part of the skeleton. Instead, Hox genes pattern the lateral plate mesoderm in a non-collinear way, independently from the somite-derived part of the skeleton (McIntyre *et al.*, 2007, Wellik, 2007).

In the past it was suggested that different combinations of Hox gene expression would define the formation of diverse structures along the AP axis – the "Hox code" (Kessel *et al.*, 1991). Later experiments showed that, even though the mutation of adjacent Hox paralogous groups resulted in partly overlapping affected areas, the morphological effects were quite different. It was then concluded that the "Hox code" in the somitic mesoderm was the result of the distinct contributions of each group expressed in a given region (McIntyre *et al.*, 2007, Wellik, 2007).

Initially, loss-of-function experiments resulted in minor effects on the axial skeleton since only part of the paralog group was mutated. These results were not in agreement with the role Hox genes supposedly had in vertebrate axial skeleton segmentation and differentiation along the AP axis. In later studies, it was concluded that Hox genes were largely redundant within each paralog group (Wellik *et al.*, 2003). *Hoxa3* and *Hoxd3* individual mutations, for example, show distinct phenotypes. However, when both genes were mutated, a synergistic effect was observed and, when the two genes were swapped, they successfully replaced each other functionally (Greer et al., 2000). For this reason, it was necessary to mutate an entire paralog group to accurately determine its importance in vertebrate axial patterning. The loss of function of *Hox10* genes causes the formation of rib-bearing vertebrae in the place of lumbar vertebrae (that normally do not have ribs) that extend past the sacral region (Fig 6). A mechanism arose by which the vertebrae have ribs as a ground-state from head to tail that is repressed in some regions of evolved vertebrates. According to this hypothesis, *Hox10* genes would repress rib formation in the lumbar and sacral regions. Since the knock-out of paralogous group 11 causes sacral vertebrae to assume a lumbar identity instead, *Hox11* genes were proposed to partially repress *Hox10* activity (Wellik *et al.*, 2003).

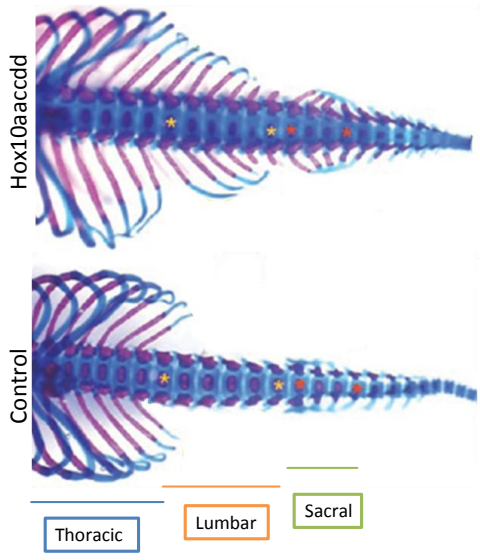


Figure 6 Effect of *Hox10* loss-of-function in the axial skeleton.
Axial skeletons of a *Hox10* triple mutant (top) and a control at E18.5 (bottom). Vertebra types are identified. Adapted from Wellik and Capecchi, 2003.

Several loss-of-function experiments demonstrated that the functional domains of Hox genes did not always correlate with their previously described boundary of expression (Carapuço et al., 2005). Indeed, *Hox10* genes were shown to be functionally important up to the thoracic/lumbar transition, where its anterior expression boundary was expected to be observed (Wellik *et al.*, 2003). However, the expression domain of these genes was reported to rarely extend that far (Burke et al., 1995). It was later shown that the *Hox10* expression domain is not stable through the entire embryonic development. The anterior expression border of the three *Hox10* genes corresponds to the functional domain only while somites at the thoracic/lumbar transition are being formed. In later stages of development, the domain of expression becomes progressively more posterior. In order to assess if *Hox10* genes were functionally relevant at the stage of somitic formation, the Delta-like1 (Dll1) promoter was used to drive the expression of the *Hoxa10* gene (Carapuço et al., 2005). The mouse Dll1 is a homologue of the *Drosophila* Delta gene. It is expressed in presomitic mesoderm (PSM) as well as in newly formed somites (Bettenhausen *et al.*, 1995, Beckers *et al.*, 2000). The overexpression of *Hoxa10* in the PSM resulted in striking skeleton abnormalities. The highest transgene copy number caused the removal of all thoracic vertebrae (Carapuço et al., 2005). This observation is consistent with the rib-repressing function attributed to the *Hox10* paralogous group through loss-of-function experiments (Wellik *et al.*, 2003). The ribless phenotype was also characterized by the presence of an ossified sternum with no sternebrae (segments of the sternum), larger cervical vertebrae than normal and the absence of the cartilaginous fusions that usually form in the sacral region of the skeleton (Fig 7). Low-copy transgenes showed milder defects, affecting just the first and two last rib-bearing vertebrae. *Hox10* genes were therefore demontrated to be functionally relevant in axial skeleton patterning at the presomitic mesoderm level, before the formation of somites. Aside from *Hox10* group genes, *Hox11* and *Hox6* paralogous groups also produced mutant phenotypes with the use of the same experimental approach. Dll-Hoxb6 embryos, in agreement with the knock-out experimental approaches, resulted in the formation of rib-bearing vertebrae from the second cervical vertebrae (C2) to the caudal region (Mallo lab, unpublished data). In fact, this paralog group was reported to be essential in rib cage patterning (McIntyre et al., 2007).
Hoxa11 overexpression was expected to expand sacral and caudal morphologies to more anterior parts of the skeleton. Indeed, Dll-Hoxa11 embryos showed fused ribs (a sign of sacralization) and a more anterior position for the sacrum. However, the use of a promoter of a gene expressed in formed somites driving *Hoxa11* also showed a mutant phenotype. This means that, although Hox10 and Hox6 proteins regulate the necessary genes for axial skeleton patterning before or just after somite formation, Hox11 proteins seem to need signals from both the pre-somitic and somitic mesoderm to fully perform its function.

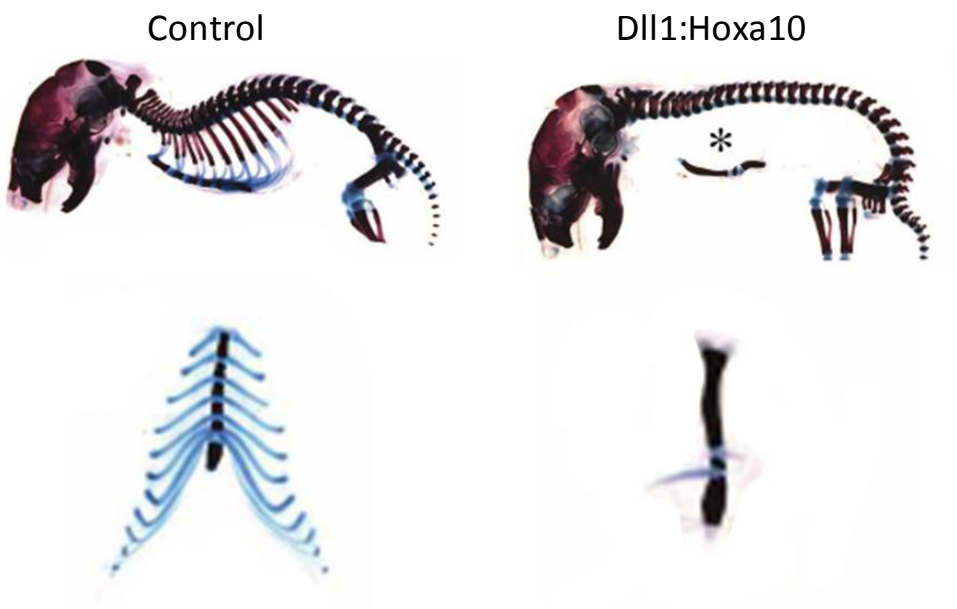


Figure 7 Patterning activity of *Hoxa10* overexpressed in the PSM.
At the top a general view of a wild-type (left) and a transgenic (right) skeleton is showed. The asterisk indicates the absence of thoracic ribs. At the bottom an anterior view of the sternum and the associated cartilaginous part of the ribs is showed for both wild-type (left) and transgenic (right) embryos. Adapted from Carapuço et al., 2005.

I.4 Objective

Hox10 genes were shown to have a rib-repressing activity both by loss-of-function and overexpression experiments (Wellik *et al.*, 2003, Carapuço *et al.*, 2005). The aim of this work is to understand what amino acid residues are necessary for *Hox10* genes to achieve their functional specificity in the axial skeleton. An octapeptide located just N-terminal to the HD, conserved in all the members of the paralog group 10 and absent in all other Hox proteins, has been identified. Four constructs were generated and overexpressed in the PSM to determine if this peptide motif is necessary or even sufficient to repress the formation of ribs in thoracic vertebrae. This study could extend the present knowledge on how *Hox* genes achieve such precise functional specificity *in vivo* that contrasts with their apparent lack of DNA-binding specificity *in vitro*.

II. Material and methods

II.1 Mice strain and housing conditions

The animal model used for this study was the mouse. All animals were from the FVB strain maintained on a 12-h dark/light cycle in a specific pathogen–free animal facility at *Instituto Gulbenkian de Ciência* (IGC).

II.2 Making of transgenic constructs

Transgenic constructs were generated using standard molecular biology techniques.

II.2.1 Mutagenesis by PCR

Most transgenic constructs were generated using a mutagenesis protocol based on PCR (polymerase chain reaction) (Fig 8). Four primers were used to generate each of the mutated and chimeric constructs: two of them annealed perfectly to each of the extremities of the DNA fragment and the other two were internal, overlapping primers encompassing the area to be modified or the border area in chimeric constructs. The primers used are detailed in table 1 and the PCR conditions used are shown in table 2. The Pfu DNA polymerase was used for its proof-reading abilities.

For each construct, two separate PCR reactions were initially performed. Both used one of the extremity primers and one of the internal primers, generating two complementary fragments. The resulting PCR products were ran in an agarose gel and the bands with the expected size were cut out and purified with the QiaexII Gel Extraction kit. Then, equimolecular amounts of the two fragments were mixed in a single reaction which had all the standard PCR components except for primers. To specifically amplify the few fragments that span the whole region of interest and contain the desired mutation, both extremity primers were then added to the reaction. After the final reaction was completed the DNA was purified by phenol-chloroform extraction. For this, TE buffer was used to make a final volume of 100µL and an equal volume of phenol-chloform was added. The samples were mixed and centrifuged for 4min at 14000 rpm. The DNA was recovered from the aqueous phase and precipitated with 0.1 of 3M sodium acetate (pH5.2) and 2.5 volumes of 100% ethanol for 30min at -80ºC. The precipitated DNA was recovered by centrifugation at 14000rpm for 10min at 4ºC. The supernatant was discarded and the DNA pellet was air-dried. The precipitaded DNA was resuspended in an appropriate volume of water for further cloning.

The Hoxa10ΔOct construct which lacked the octapeptide had been previously generated. For this reason the mutated *Hoxa10* sequence was amplified using the same conditions as the final PCR reaction step (Table 2 - bottom).

GENERATION OF CHIMERIC PROTEINS

An adapted version of the initial mutagenesis by PCR protocol was also used to generate chimeric proteins. Two kinds of chimeric proteins were produced. In the case of the Hoxb9Oct construct, only a small peptide motif was swapped by that of a different protein (Fig 8b). In this case, the internal primers were designed to flank the part of the sequence that will be swapped and include the sequence to be inserted as a 3'overlapping extension. However,

there was a second type of chimeric protein generated composed partly of Hoxa10 and partly of Hoxb9 (Hoxa10b9). In this case, two different templates were used in each of the initial PCR reactions (Fig 8c).

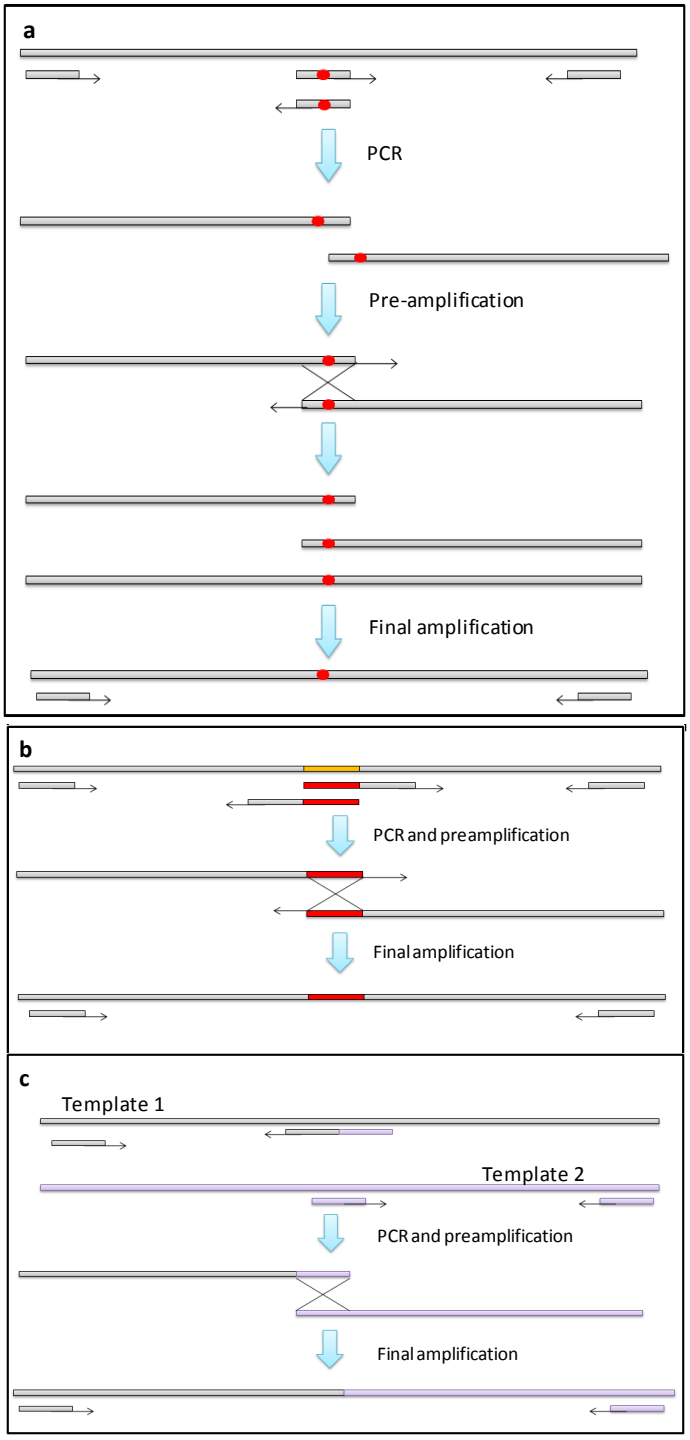


Figure 8 Schematic representation of the standard mutagenesis by PCR protocol.
a) standard procedure – the red dot indicates the inserted mutation. b) and c) generation of chimeric proteins to a different extent. In b) the inserted sequence is marked in red. In c) two different templates (illustrated with different colors) are used.

Table 1 Sequences of the primers used for mutagenesis. The nucleotides that do not anneal to the template sequence are marked in red.

	Internal primers	
	Orientation	Sequence
ΔOct		*
		*
TSAA	Forward	5'AACTGGCTCGCAGCAAAGGCCGGCCGGAAGAAACG3'
	Reverse	5'TTCTTCCGGCCGGCCTTTGCTGCGAGCCAGTTG3'
B9Oct	Forward	5'AACTGGCTCACAGCAAAGAGCGGCCGGAAAAAGCGCTGTCCCTAC3'
	Reverse	5'GCCGCTCTTTGCTGTGAGCCAGTTGGCGGAGGGGTTGGTTTGATC3'
A10-B9	Forward	5'GTACCCGGCTACTTCCGCCTGTCGGTCCCGCCGGCCGAGAGCAGG3'
	Reverse	5'CGACAGGCGGAAGTAGCCGGGTACCGGGACGCCCTGGGGCTGGATG3'

	External primers		
	Orientation	Sequence	Restriction site
ΔOct	Forward	5'CTGGATCCTGCTCGGAGAGCCCTGCCGCG3'	BamH1
	Reverse	5'CAGCGGCCGCCGGCACAGGTGTGAGTTCTG3'	Not1
TSAA	Forward	5'GCCTGCAGGCCTACGGCACGGCC3'	Pst1**
	Reverse	5'CAGCGGCCGCCGGCACAGGTGTGAGTTCTG3'	Not1
B9Oct	Forward	5'CGGGATCCATTTCTGGGACGCTTAGC3'	BamH1
	Reverse	5'CGGCGGCCGCAGTCGTCACATAACTAAGAG3'	Not1
A10-B9	Forward	5'GCCTGCAGGCCTACGGCACGGCC3'	Pst1**
	Reverse	5'CAGCGGCCGCCGGCACAGGTGTGAGTTCTG3'	Not1

* This construct did not require the use of internal primers.
** Although these restriction sites were in the primer sequence, this enzyme was never used for the subsequent cloning steps; a HindIII restriction site in the fragment was used instead.

Table 2 PCR conditions for mutagenesis

Initial PCR	Primer 1	1µL	**Temperature**	**Time (minutes)**	
	Primer 2	1µL	95ºC	4	
	Template (10ng/µL)	1µL	95ºC	1	
	10x buffer (+20mM $MgSO_4$)	5µL (1x)	62ºC	1	30 cycles
	dNTPs (25mM)	1µL	72ºC	1.5	
	Pfu (2.5U/µL)	1µL	72ºC	7	
	H_2O	to 50µL			

Preamplification	Fragment 1	*	**Temperature**	**Time (minutes)**	
	Fragment 2	*	95ºC	4	
	10x buffer (+20mM $MgSO_4$)	3µL (1x)	95ºC	1	
	dNTPs (25mM)	0.6µL	62ºC	1	4 cycles
	Pfu (2.5U/µL)	1µL	72ºC	1.5	
	H_2O	to 30µL	72ºC	7	

* equimolecular amounts were used

Final amplification	Primer 1 (25mM)	1µL	**Temperature**	**Time (minutes)**	
	Primer 2 (25mM)	1µL	95ºC	4	
			95ºC	1.5	
			62ºC	1.5	35 cycles
			72ºC	2	
			72ºC	7	

II.2.2 Molecular cloning

The mutated, chimeric and unaltered PCR fragments obtained were cloned downstream of a DNA sequence that coded for the FLAG-tag, placing them in the same open reading frame. To obtain transgenic mice the resulting constructs were cloned downstream of the Dll1 promoter and upstream of the *Hoxa10* 3'UTR and the polyA tail. The pBluescript® II KS+ phagemid was used as a vector for these cloning steps. However, for *in vitro* protein production, the same constructs were cloned into the multiple cloning site of the pCMV-Sport6.1 vector.

DIGESTION WITH RESTRICTION ENZYMES AND LIGATION

After obtaining the mutated sequences these were digested with the appropriate restriction enzymes (see restriction sites in Table 1). Either way, the fragment of interest was ligated to a vector attached to the sequence that encodes the FLAG-tag. The amount of DNA used in digestions was usually 3µg. The total volume of the digestion reaction was 20µL. The completion of the digestion was verified by agarose gel electrophoresis. The whole reaction was then loaded on an agarose gel and the bands of interest were then cut out and purified using the QIAEX II Gel Extraction kit. An equimolar amount of vector and insert, 1U of the T4 DNA ligase, and 1x ligation buffer were then used in a ligation reaction. At all times a negative control was employed, using the same conditions except for the absence of insert DNA. The mix was then left at room temperature or incubated at 16ºC for 2h or more. The ligation reaction was directly used to transform DH5-α competent cells as described below.

PREPARATION OF COMPETENT CELLS

The DH5-α strain of *E. coli* was used for all transformations performed in this study. Before these cells can be transformed they need to go through a process that allows the intake of exogenous DNA. The cells were inoculated from a frozen stock in 1mL of LB and grown for 16-20h at 37°C, with shaking (225rpm). Then, 1mL of the culture was diluted into 100mL of fresh LB and incubated at 37°C, 225rpm, until the culture reached a 600nm optic density of 0.5. The culture was chilled on ice for 15min and the bacteria collected by centrifugation at 4000rpm for 15min at 4ºC. The supernatant was removed and the pellet was resuspended in 33mL of ice-cold RF1 buffer (1/3 of the volume collected) and incubated on ice for 1h. The resulting solution was centrifuged with the same conditions as before and the supernatant discarded. The cell pellet was then resuspended in 8mL of pre-chilled RF2 buffer (1/12.5 of the volume collected) and left on ice for 15min. The final volume was aliquoted into 1.5ml centrifuge tubes, flash-frozen, using a bath of dry ice and 100% ethanol. The aliquots were stored at -80°C.

TRANSFORMATION OF DH5-α CELLS

3µL of the ligation reaction were added to 50µL of DH5-α competent cells and incubated on ice for 25min. The bacteria were incubated at 42ºC for 45 seconds and chilled on ice for 2min. After adding 1mL of LB medium, the cells were incubated at 37ºC for 45min with shaking. The culture was centrifuged at 4000rpm for 4min and about 800µL of the supernatant medium was removed. The remaining supernatant was used to resuspend the cell pellet in order to concentrate the cell suspension. The bacteria were then plated on solid LB medium with ampicillin (50µg/mL) and incubated overnight at 37ºC.

PLASMID DNA MINI-PREPARATION

Single colonies were picked using pipette tips and grown on 3mL LB medium with ampicillin (50µg/mL) at 37°C with shaking overnight. 1.5mL of the culture were centrifuged at 4000rpm for 4min and the pellet resuspended in 100µl of TE with RNase (10µg/mL). Then, 300µl of TENS were added, the mix was vortexed and 150µl of 3M potassium acetate (pH5.2) were added. After this, the mixture was centrifuged 4min at 14000rpm. The supernatant was transferred into a fresh tube that contained 900µl of 100% ethanol. After mixing and centrifuging 4min at 14000rpm, the supernatant was removed and the pellet air-dried and resuspended in 50µl of TE. The resulting plasmid preparation was then screened for positive colonies by digestion with the appropriate restriction enzymes.

When higher purity DNA was required for experiments such as sequencing reactions, plasmids were purified using plasmid preparation kits. Briefly, the protocol is based on plasmidic DNA binding to an anion-exchange resin and subsequent elution after the cell lysis and plasmidic DNA precipitation steps. When small amounts of the plasmid were needed, for example in intermediate cloning steps, a mini-preparation was conducted using a QIAprep® spin miniprep kit. However, when larger amounts of plasmid DNA were needed, usually in the final cloning step, a midi-preparation was conducted using the NucleoBond® plasmid DNA purification kit. In both cases, the protocol was followed according to the manufacturer specifications and the concentration of plasmidic DNA was measured using the NanoDrop® ND-1000 spectrophotometer.

II.2.3 Sequencing reaction

After the first cloning step, all sequences were checked. The primers used in the sequencing reaction annealed to RNA polymerase promoters T3 and T7, which flank the multiple cloning site of the pBluescript® II KS vector. Each 10µL reaction contained 2µL of BigDye® terminator sequencing buffer (5X), 2µL BigDye® terminator ready reaction mix, 500ng of template DNA and 5pmol of primers. The PCR conditions were the following:

Temperature	**Time**	
96ºC	1min	
96ºC	10sec	25 cycles
50ºC	5sec	
60ºC	4min	
4ºC	until ready to purify	

After the PCR, the reaction product was transferred to a microcentrifuge and the amplified DNA was precipitated (see DNA precipitation in section II.2.1). After mixing, the tubes were incubated at room temperature for 30min and centrifuged at 14000rpm for another 30min at 4ºC. The supernatant was removed and the pellet was washed in 250µL of 70% ethanol. The tubes were centrifuged at 14000 rpm for 15min at 4ºC. The supernatant was again removed and the pellet was air-dried. The samples were then sent to the IGC sequencing service. The resulting output sequences were analyzed in detail by the combined use of the BLAST (http://www.ncbi.nlm.nih.gov/BLAST/), Bioedit, Chromas and Sequence Analysis software.

II.2.4 Agarose gel electrophoresis

This technique that is used to separate and visualize nucleic acids was necessary to check the efficiency of restriction enzyme digestions, to purify specific DNA fragments and to check the integrity of a DNA sample, among other applications. Agarose was dissolved in 1X TAE, usually at a concentration of 0.8%. However, in some cases a 2% agarose gel was used instead. Ethidium bromide was added in order to visualize the DNA with the use of a UV light to a final concentration of 0.2µg/mL. Loading buffer was added to each sample to a 1x final concentration and a DNA ladder was used to estimate the size of the DNA fragments. An electric current of 100-120V was applied to the gel immersed in 1X TAE buffer.

BAND PURIFICATION

DNA fragments were purified from an agarose gel, using the QIAEXII gel extraction kit, according to the protocol recommended by the manufacturer. In the final step, the DNA fragments were eluted with TE.

The QIAquick gel extraction kit was used to prepare DNA fragments for microinjection, according with the protocol recommended by the manufacturer. The concentration of the DNA construct was measured using the NanoDrop® ND-1000 spectrophotometer.

II.3 Microinjection

Transgenic embryos were generated by pronuclear injection, which was carried out by personal at the IGC's transgenics unit. Briefly, female mice were superovulated with the use of hormones and mated with males. The fertilized oocytes were then recovered and DNA fragments (Fig 9) were microinjected into one of the pronuclei, which are subsequently implanted in the oviduct of pseudopregnant females.

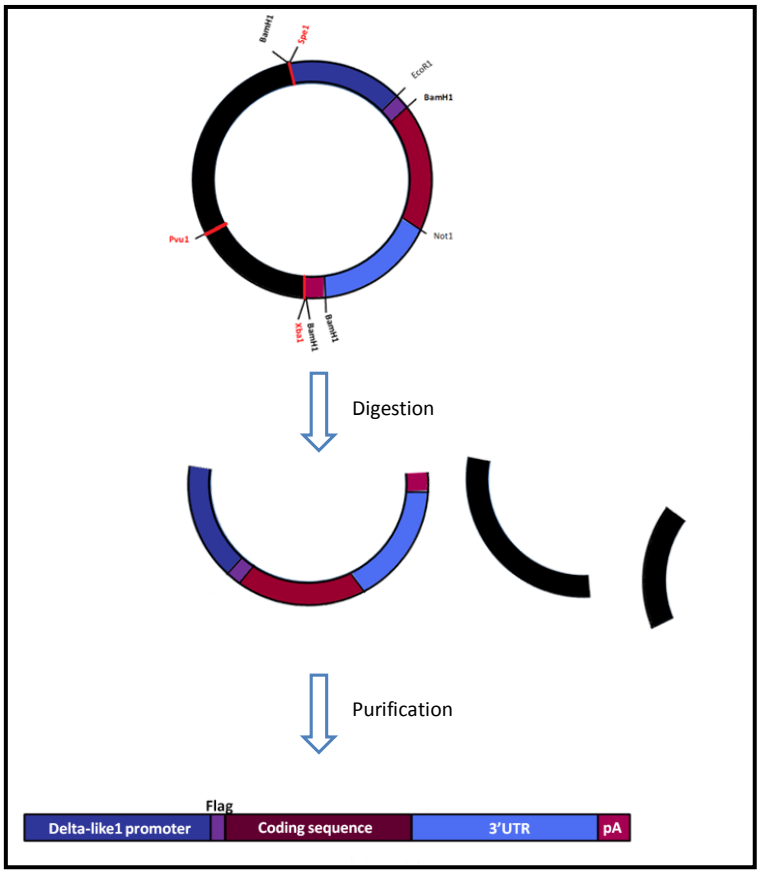


Figure 9 Final cloning step.
The final construct contains the Delta-like1 promoter, an N-terminal FLAG-Tag, the mutated coding sequence, the *Hoxa10* 3'UTR and the polyA tail (pA).

II.4 Embryo collection

All embryos were collected at embryonic day (E) 18.5 by caesarean section.

II.4.1 Embryo genotyping

To identify transgenic embryos, the intestine of each embryo was used for genotyping. The intestines were digested in 500µL of Laird's buffer with 100µg/mL of proteinase K at 50ºC, overnight, with shaking. In order to precipitate the genomic DNA, 500µL of isopropanol were added and mixed by gently inverting the tubes. The DNA was "fished" with a pipette tip and dissolved in 250µL of TE buffer at 37ºC with shaking for 3h or more. The presence of the transgene was determined by PCR using the conditions shown in table 4. When mouse cDNAs were used in the transgenic constructs, the primers were designed to flank an intron to distinguish it from the endogenous DNA. A bigger band will be the result of endogenous genomic DNA amplification and a smaller band will be the consequence of the presence of the microinjected cDNA. For chimeric constructs, it was sufficient to use one primer that annealed to the cDNA of one of the proteins and a second primer that annealed to the cDNA of the other protein. The sequences of the primers designed to genotype each construct is shown in table 3.

Table 3 Sequences of the primers used for genotyping.

	Orientation	Sequence
DllHoxa10EFGpA	Forward	5'AGCGAGTCCTAGACTCC3'
DllHoxa10TSAApA	Reverse	5'GTCCGTGAGGTGGACGCTACG3'
DllHoxb9Oct-A103'UTRpA	Forward	5'GTCACGAGAGTGAGGACGCGC3'
	Reverse	5'CGGCGGCCGCAGTCGTCACATAACTAAGAG3'
DllHoxba10/b9-A103'UTRpA	Forward	5'AGCGAGTCCTAGACTCC3'
	Reverse	5'CGGCGGCCGCAGTCGTCACATAACTAAGAG3'

Table 4 Conditions for genotyping PCR reactions.

DNA	1µL
10x buffer	2,5µL (1x)
DMSO	2µL (8%v/v)
Primer 1 (25mM)	0,25µL
Primer 2 (25mM)	0,25µL
dNTPs (25mM)	0,2µL
Taq (5U/ µL)	0,2µL (1U)
H_2O	to 25µL

Temperature	Time	
95ºC	4 min	
95ºC	45 sec	
65ºC	45 sec	35 cycles
72ºC	45 sec	
72ºC	7 min	

II.4.2 Skeletal analysis

Embryos collected at E18.5 were eviscerated, skinned and fixed in absolute ethanol. In order to stain the cartilage part of the skeleton, the fetuses were incubated in an alcian blue solution for 12-20h at room temperature and then incubated in 100% ethanol overnight. The tissues were then partially digested with 2% Potassium Hydroxide and the bones were stained with an alizarin red solution. To complete soft tissue digestion, the embryos were further digested in 2% Potassium Hydroxide for 12-20h and stored in 25% glycerol.

II.5 *in vitro* protein analysis

II.5.1 *in vitro* protein synthesis and detection

All wild-type, mutant and chimeric proteins were synthesized by *in vitro* translation using the TNT® Sp6 reticulocyte lysate system. This system allows for both transcription and translation reactions to take place in a single experimental step. The protocol was followed according to the manufacturer instructions. During the reaction, biotinylated lysine residues were incorporated into the proteins to allow their detection using the Transcend™ Non-Radioactive Translation Detection System.

In order to do this, 1µL of each TNT reaction was added to 15µL of SDS-containing loading buffer and incubated at 70ºC for 15min. The samples were loaded on a gel composed by a 5% SDS polyacrylamide stacking gel and a 10-12% SDS polyacrylamide resolving gel. They were then run in running buffer at 110V, for about 1.5h. To estimate protein sizes an appropriate protein ladder was used. After the run, the proteins were transferred to a PVDF membrane previously wetted in methanol, washed in water and equilibrated in transfer buffer. In order to

do this, the "sandwich" represented in figure 10 was assembled and immersed in transfer buffer at 4ºC. A 200mA electric current was applied for 1 hour. The proteins were visualized in the polyvinylidene difluoride (PVDF) membrane by Streptavidin-Alkaline Phosphatase (#V5591) from Promega that binds to the biotinylated lysines added to the *in vitro* synthesized proteins. The protocol was followed according to the manufacturer specifications.

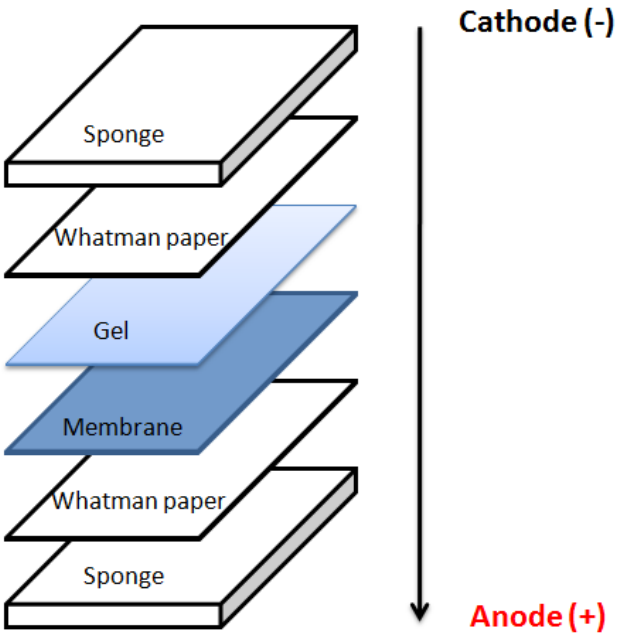


Figure 10 Protein transfer schematic

II.5.2 Electrophoretic mobility shift assay

The electrophoretic mobility shift assay (EMSA) was used to verify if the modified proteins had the ability to bind to a specific target DNA. The probe was synthesized by PCR amplification of approximately 100bp of the promoter region of the *Lbx1* gene, using the primers and the annealing temperature described in table 6. The remaining PCR conditions were the same that were used for genotyping, shown in table 4. The reaction product was then purified with the QIAquick PCR purification kit according with the manufacturer specifications. 5µL of the fragment were labeled with ^{32}P-ATPγ, using 10-20U of T4 kinase and 1X T4 kinase buffer in a 20µL final volume. The reaction was incubated for 2h at 37ºC. The labeled probes were purified using the QIAquick PCR purification kit and stored at -20ºC. 3µL of reticulocyte lysate were incubated with a variable amount of DNA probe, 2x binding buffer and 1 µg of poly-dIdC (a non-specific competitor) in final volume of 20 µL. After incubating for 30min at room temperature, the samples were run in a 5% non-denaturing polyacrylamide gel at 150V, in 0.5x TBE. The gel was previously run for 30-60min at 100V. After the run, the plates were disassembled and the gel was placed on top of a precut 3mm Whatman paper. The gel was then covered with a plastic wrap and dried in a vacuum heating device at 80ºC for 1 hour. The covered dry gel was exposed to an X-ray film overnight at -80ºC. If a weak signal was observed, the film was exposed for longer time periods.

Table 5 Primer sequences for *Lbx1* probe synthesis

	Orientation	Sequence	Ta
Lbx	Forward	5'CAAACATCTCGGGCCGAGCAC3'	61ºC
	Reverse	5'TCTGTCCAATGCAGGCGGCTC3'	

II.6 Cell culture

The cells used for transfection were C3H-10T1/2 cells and 293-T cells. Both lines of cells were grown in Dulbecco's Modified Eagle's Medium (DMEM) containing 20% Fetal calf serum (FCS), 1x penicillin and streptomycin and 1x L-glutamine. After thawing the cells in a 37ºC water bath, they were transferred to 5mL of growth medium, centrifuged at 1000rpm for 5min and the supernatant was discarded. Cells were always plated using a media volume to surface area ratio of 0.2-0.3ml/cm^2. The pellet was therefore resuspended in an appropriate amount of media, plated and incubated overnight at 37ºC, in a humidified environment containing 5%CO_2. The medium was changed daily until cells reached confluency, at which point it was necessary to pass the cells. For this, the media was removed and the cells were washed with Dulbecco's phosphate buffered saline (PBS). In order to detach the cells, 30-35µL/cm^2 of trypsin with EDTA were added and the plates were incubated at 37ºC for about 5min. When cells detached, trypsin was inactivated by addition of about 10 times the trypsin volume of complete medium. Cells were then recovered by centrifugation at 1000rpm for 5min. The cells were then resuspended in an adequate volume of growth media, plated and incubated at 37ºC, in a humidified environment containing 5%CO_2.

II.6.1 Transfection

The media was changed daily until cells became 90-95% confluent and were ready for transfection. About three hours prior to transfection, the medium was changed for a similar one without antibiotics. Transfections were performed using Lipofectamin™ 2000 according to manufacturer instructions. 8µg of the pCMVSport6.1 containing both the unaltered *Hoxa10* cDNA (Sp6.1Hoxa10) and the *Hoxa10* with two potential phosphorylation sites mutated (Sp6.1Hoxa10TSAA) were transfected (Table 5). A GFP-expressing plasmid was also used to control for transfection efficiency.

Table 6 List of trasnfected samples and negative controls

cell line	**construct**
293T	none
293T	Sp6.1Hoxa10
293T	Sp6.1Hoxa10TSAA
10T1/2	none
10T1/2	Sp6.1Hoxa10
10T1/2	Sp6.1Hoxa10TSAA

II.6.2 Cell lysis

After transfection, cells were incubated for about 24h. At this point, cells were trypsinized again as described above. They were then washed twice in ice-cold Dulbecco's PBS by gentle resuspension and collected by centrifugation at 1000rpm, for 5min, at 4ºC. Cells were lysed by resuspension in 100-400µL of an SDS/NP-40 buffer followed by a 15min incubation on ice. After centrifugation at 14000rpm, for 10min, at 4ºC, the protein-containing supernatant was transferred to a new tube and stored at -80ºC. The pellet was discarded.

II.6.3 Cell extract protein analysis

WESTERN BLOT

The western blot procedure was used to visualize the production of Hoxa10 and Hoxa10TSAA proteins in C3H-10T1/2 cells and 293-T cells. 5µL of 293-T samples and 10-15µL of C3H-10T1/2 samples were added to an equal amount of SDS-containing loading buffer and incubated for about 3min at 100ºC. The SDS polyacrylamide gel electrophoresis (PAGE) and protein transfer was conducted as described in section II.4.1. However, a chemiluminescent detection method was used to detect the proteins. After transfer, nonspecific binding sites were blocked by incubation with a blocking buffer for 1h with shacking at room temperature. The monoclonal primary antibody ANTI-FLAG® M2 (#F1804) from Sigma-Aldrich raised in mouse was used in a 1:1000 dilution in blocking buffer. The membrane was incubated in the antibody solution at 4ºC overnight with shaking. The blot was then washed 2-3 times with PBS with 0.1% Tween20 (PBT) for 10min and incubated with a goat α-mouse secondary antibody conjugated with the horseradish peroxidase (HRP) enzyme diluted 1:2000 in blocking buffer, for 1h at room temperature. The membrane was washed in PBT 3 times for 15min. The proteins were detected using the Pierce SuperSignal® West Pico Chemiluminescent HRP substrate and the protocol employed followed the recommended procedure. A film was exposed for variable periods of time and developed manually.

The underlined solutions are detailed in *Appendix I – Buffers, Solutions and Media*

III. Results

In order to facilitate detection of the Hoxa10 protein derivatives generated, a FLAG-tag was added N-terminally to each of them. To ensure that the presence of the FLAG-tag was not influencing protein function, a construct containing the *Hoxa10* cDNA downstream of the sequence coding for the FLAG-tag was overexpressed in the PSM. Two phenotypes with different intensities were obtained (Fig 11). One of the phenotypes was completely ribless, while the other only lacked ribs in the T1, T12 and T13 thoracic vertebrae. These phenotypes are consistent with the overexpression of *Hoxa10* cDNA without the FLAG-Tag attached (Carapuço et al., 2005). This indicates that the use of the FLAG-Tag does not influence the rib-repressing activity of the protein when overexpressed in the PSM.

To identify the motifs responsible for Hoxa10 protein functional activity, Hox protein sequences were analyzed. Two peptide motifs outside the HD of Hox10 proteins showed complete conservation in all paralogous group 10 members and were absent in all the remaining Hox proteins. Both peptides lie adjacent to the HD (Fig 12a): the NWLTAKSG octapeptide is located upstream of the HD, while the RENRIRELT peptide is located C-terminally to it. To determine their contribution to the rib-repressing function of Hox10 proteins, the peptide motifs were removed and the resulting mutant proteins were overexpressed in the presomitic mesoderm driven by the Dll1 promoter. Although a construct was produced in which the RENRIRELT peptide was removed from the original protein, so far no transgenic embryos could be produced. For this reason this peptide motif will not be approached in more detail in this work. However, the octapeptide deletion resulted in interesting phenotypes that might clarify the importance of sequences outside the HD in specific functions of Hox genes.

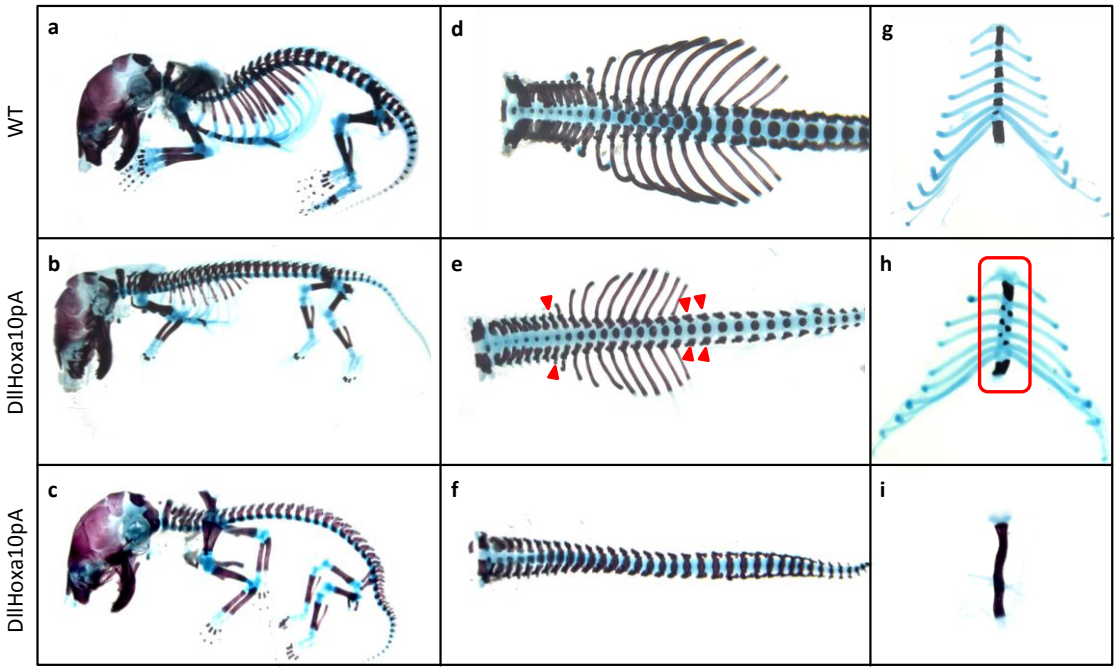


Figure 11 Skeletal staining of E18.5 embryo collected after microinjection with the DllHoxa10pA construct.
The *Hoxa10* construct is expressed with the Dll1 promoter and a FLAG-Tag attached N-terminally. A wild-type embryo (top), a mild-phenotype transgenic (middle) and a ribless transgenic (bottom) are represented. a), b) and c) lateral view; d), e) and f) dorsal view of the embryos – red triangles indicate the absence of ribs. g), h) and i) ventral part of the rib cage – the red rectangle highlights abnormal attachment of the ribs to the sternum.

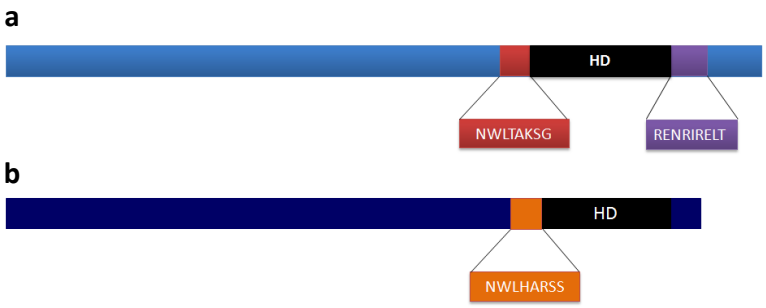


Figure 12 Schematic representation of the templates used to make the mutant constructs.
a) *Hoxa10* and b) *Hoxb9*. The Homeodomain (HD) is flanked by the conserved peptide motifs indicated.

Figure 13 Schematic representation of the constructs generated.
a) Hoxa10 missing the octapeptide - Hoxa10ΔOct, b) Hoxa10 with a serine and a threonine mutated to alanines - Hoxa10TSAA, c) Hoxb9 with the Hoxa10 octapeptide - Hoxb9Oct, d) Hoxa10/Hoxb9 chimeric protein – Hoxa10/b9.

III.1 Is the octapeptide necessary for Hox10 functional specificity?

To check if the octapeptide is necessary for the Hoxa10 protein to achieve its functional specificity, a construct of the Hoxa10 protein lacking the octapeptide was produced and overexpressed in the presomitic mesoderm (DllHoxa10ΔOctpA – Fig 13a). None of the four transgenics analyzed at E18.5 showed severe skeletal abnormalities (Fig 14). One of the embryos was missing fully-formed ribs in the first thoracic vertebra (Fig 14d) and showed abnormal attachment of the ribs to the sternum (Fig 14h). The other three transgenic embryos had normal axial skeleton phenotypes. Since *Hoxa10* overexpression in the PSM has been shown to produce a ribless phenotype (Carapuço et al., 2005), these results suggest that the octapeptide is necessary for Hox group 10 genes to achieve their functional specificity in the axial skeleton.

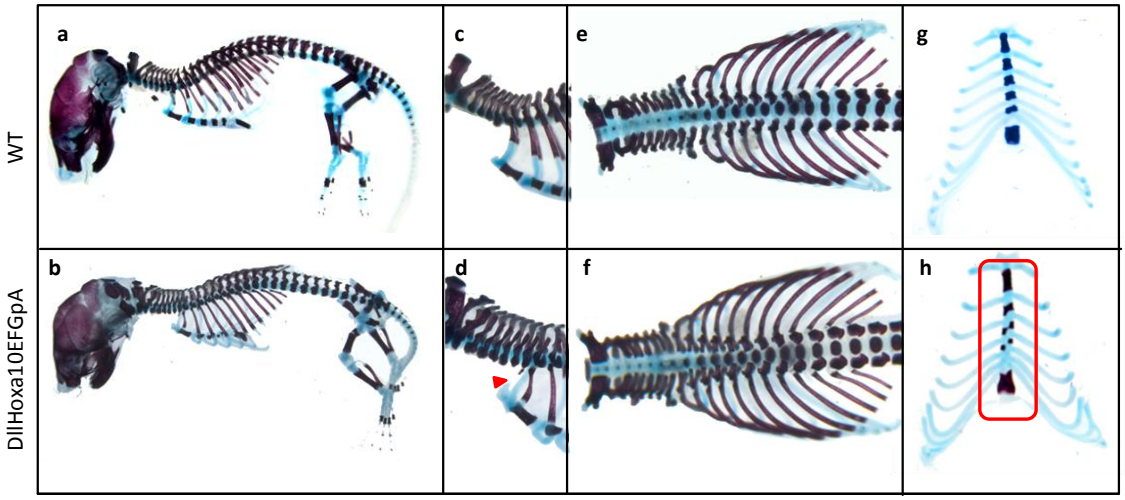


Figure 14 Skeletal staining of E18.5 embryos collected after microinjection with the DllHoxa10EFGpA construct.
The *Hoxa10* construct without the octapeptide is expressed. A wild-type embryo (top) and a transgenic (bottom) are represented. a) and b) lateral view; c) and d) close-up of the cervical and first thoracic ribs - the transgenic is missing the first rib (red triangle); e) and f) dorsal view of the embryos. g) and h) ventral part of the rib cage – the red rectangle highlights abnormal attachment of the ribs to the sternum.

III.2 The mutation of possible phosphorylation sites within the homeodomain

Since the octapeptide seems to play such an important role in Hox group 10 protein function, specific aminoacids within this peptide motif were examined more closely. Two possible phosphorylation sites, a serine and a threonine were identified. Both amino acids were mutated to alanines and the modified protein was overexpressed in the PSM (DllHoxa10TSAApA – Fig 13b). Nine transgenic embryos were obtained and none of them showed any kind of defects in the axial skeleton (Fig 15). These results strongly suggest that the mutation of these amino acids renders the protein unable to perform its function.

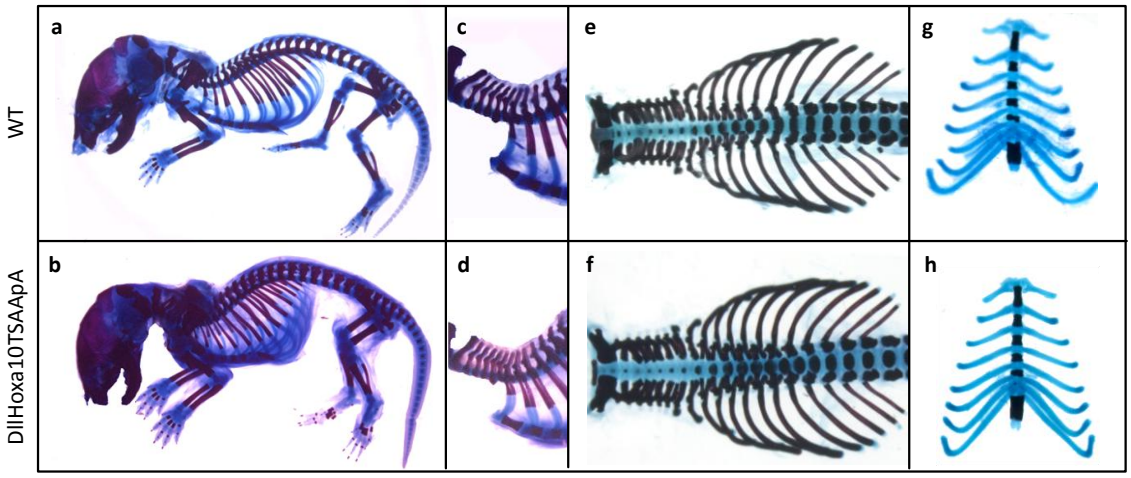


Figure 15 Skeletal staining of E18,5 embryos collected after microinjection with the DllHoxa10TSAApA.
Hoxa10 construct with two possible phosphorylation sites mutated is expressed. A wild-type embryo (top) and a transgenic (bottom) are represented. a) and b) lateral view; c) and d) close-up of the cervical and first thoracic ribs; e) and f) dorsal view of the embryos. g) and h) ventral part of the rib cage.

Since the mutated aminoacids are possible phosphorylation sites, the results raised the possibility that their phosphorylation was necessary for the Hoxa10 rib-repressing activity. In order to determine if these residues were phosphorylated *in vivo*, cells were transfected and the first steps were taken to conduct a two-dimensional polyacrylamide gel electrophoresis (2D-PAGE). Two different cell lines were transfected both with the normal *Hoxa10* (Sp6.1Hoxa10) and the construct with the two possible phosphorylation sites mutated (Sp6.1Hoxa10TSAA). It was expected that this comparison would provide initial hints as to whether Hoxa10 becomes phosphorylated in the mutated residues. 293-T cells are convenient since they have an extremely high transfection efficiency. However, C3H-10T1/2 cells are mesenchymal and might mimic more accurately the *in vivo* behaviour of presomitic mesoderm cells. In order to check if the transfected proteins were being produced by both cell lines, a western blot was conducted. The blot showed no detectable differences between mutated and non-mutated protein production. However, there was a significant band size difference between 293-T and C3H-10T1/2 samples: protein bands from 293-T lysates had the expected size (about 41KDa) but C3H-10T1/2 produced smaller proteins from the same initial cDNA (Fig 16). The meaning of this finding is still unknown but it will be further explored in future experiments. Ultimately, these samples will be run in a 2D-PAGE that has not been conducted yet due to time constraints.

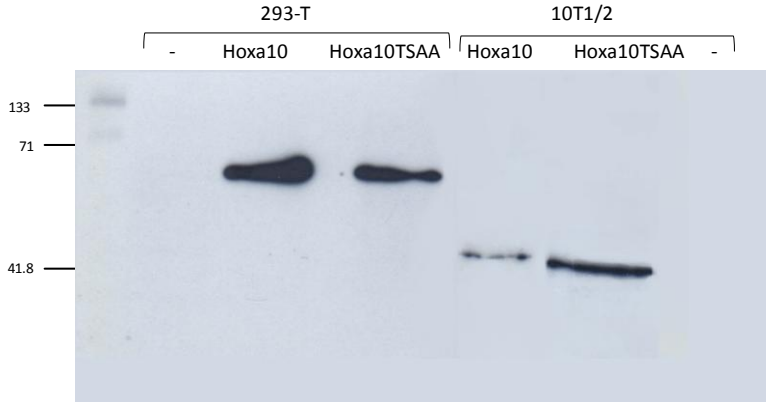


Figure 16 Western blot for transfected 293-T and 10T1/2 cell type lysates with Hoxa10 and Hoxa10TSAA.
The Hoxa10TSAA has a threonine and a serine mutated to alanines. There is also a negative control where no DNA was transfected. Protein molecular weight indicators are in KDa.

III.3 Is the octapeptide sufficient to produce a ribless phenotype?

The approach taken to determine if the octapeptide is sufficient for Hoxa10 to achieve its rib-repressing activity was to select a gene which when overexpressed does not produce a phenotype in the axial skeleton phenotype and use it to generate chimeric constructs. Previous data from Mallo's lab suggested that the *Hoxb9* gene was an adequate candidate for this experiment. This gene is very closely related to *Hox10* genes and it codes for a protein that possesses a peptide motif located just upstream of the HD that differs in three amino acids from the Hox10 octapeptide (Fig 12b).
For consistency reasons, every construct involving both the presence of *Hoxa10* and *Hoxb9* cDNA possessed the *Hoxa10*'s 3'UTR, in accordance with the other constructs generated. To make sure that the phenotypes obtained would result only from the modified cDNA, the

unaltered *Hoxb9* cDNA was cloned into the plasmid containing the *Hoxa10* 3'UTR (DllHoxb9-A103'UTRpA). Three transgenics resulted from the microinjection and no deficiencies were detected in the axial skeleton (Fig 17). Therefore, the presence of the 3'UTR seems not to be enough to cause a mutant phenotype.

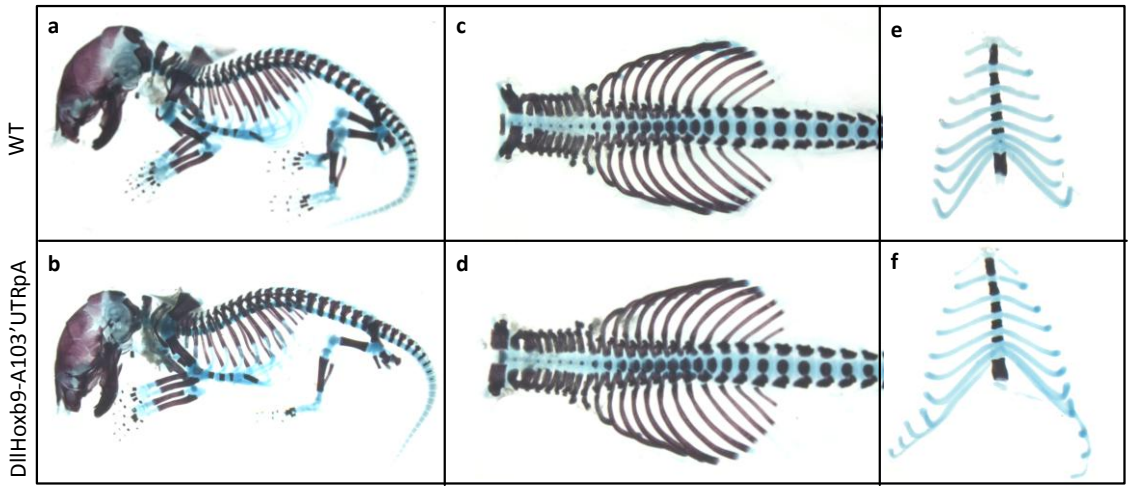


Figure 17 Skeletal staining of E18,5 embryos collected after microinjection of the DllHoxb9-A103'UTRpA construct.

The *Hoxb9* construct with the *Hoxa10* 3'UTR is expressed. A wild-type embryo (top) and a transgenic (bottom) are represented. a) and b) lateral view; c) and dorsal view of the embryos; e) and f) ventral part of the rib cage.

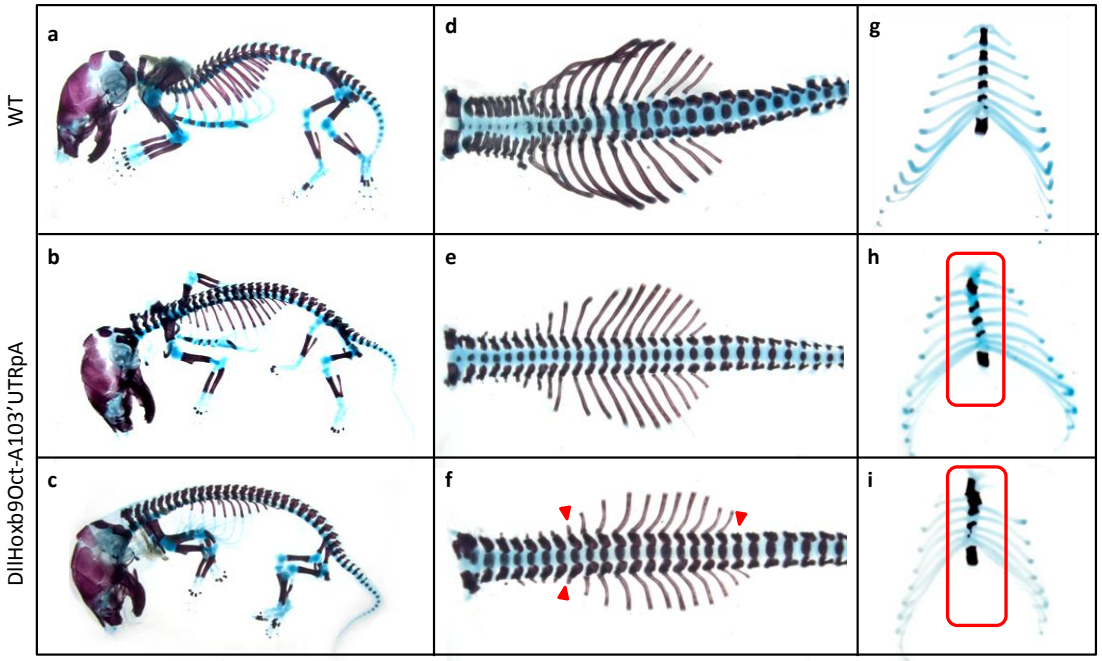


Figure 18 Skeletal staining of E18,5 embryos collected after microinjection of the DllHoxb9Oct-A103'UTRpA construct.

A *Hoxb9* construct with the *Hoxa10* octapeptide inserted just before the homeodomain is expressed. A wild-type embryo (top), a mild-phenotype transgenic (middle) and a transgenic with a more severe phenotype (bottom) are represented. a), b) and c) lateral view; d), e) and f) dorsal view of the embryos – red triangles indicate the absence of ribs. g), h) and i) ventral part of the rib cage – the red rectangle highlights abnormal attachment of the ribs to the sternum.

To determine if the Hoxa10 octapeptide was sufficient to incorporate a rib-blocking phenotype into the Hoxb9 protein the NWLHARSS motif was replaced by the Hox10 octapeptide (NWLTAKSG) (DllHoxb9Oct-A103'UTRpA – Fig 13c). The chimeric construct was then cloned into the 3'UTR of the Hoxa10 for consistency reasons. Three transgenics expressing the chimeric construct had mutant phenotypes (Fig 18). They all showed severe sternum defects, which were probably the result of abnormal rib formation (Fig 18h-i). Additionally, the cervical vertebrae in transgenic embryos seemed to have lost their typical morphology. They were larger and most of them had underdeveloped ribs (Fig 18e-f). The most severe phenotype showed absence of a few ribs both at the anterior and at the posterior borders of the thoracic region (Fig 18f). The results obtained suggest that the octapeptide has a role in rib-repressing functions but is not sufficient on its own to produce a fully ribless skeleton.

III.4 Other functionally relevant residues

The data obtained from the transgenic embryos overexpressing the Hoxb9oct construct suggested that other parts of the Hoxa10 sequence are necessary for this protein to achieve its function. For this reason, another construct was generated that contained the Hoxa10 octapeptide and the sequences N-terminal to it, as well as the Hoxb9 homeodomain and sequences C-terminal to it. Overexpression of a cDNA coding for this chimeric protein (DllHoxa10b9-A103'UTRpA - Fig 13d) in the PSM produced three transgenics, from which only one showed a detectable axial skeleton phenotype (Fig 19). Similarly to Hoxb9Oct transgenics, the embryo had sternum defects (Fig 19f). Ribs from thoracic vertebrae T1, T2 and T13 were almost absent and the other thoracic vertebrae were smaller than those observed in the wild-type skeleton (Fig 19d). Cervical vertebrae were not as affected as in Hoxb9oct transgenics. However, they were more spaced and resembled cervical vertebrae observed in Hoxa10 hypomorphic transgenics (Fig 11e). This phenotype suggests that residues in the N-terminal part of the Hoxa10 protein are not sufficient to repress thoracic rib formation.

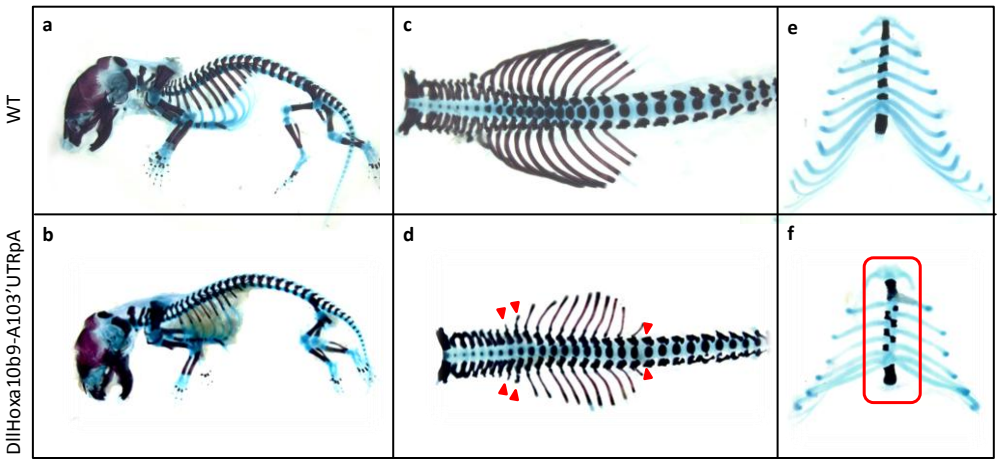


Figure 19 Skeletal staining of E18,5 embryos collected after microinjection of the DllHoxa10/b9-A103'UTRpA construct.

A *Hoxa10*/*Hoxb9* chimeric construct with the octapeptide inserted just before the homeodomain is expressed. A wild-type embryo (top) and a transgenic (bottom) are represented. a) and b) lateral view; c) and d) dorsal view of the embryos red triangles indicate the absence of ribs; e) and f) ventral part of the rib cage– the red rectangle highlights abnormal attachment of the ribs to the sternum.

III.5 Protein synthesis and DNA-binding ability

Since the sequence of the *Hoxa10* cDNA was manipulated, the mutated constructs were further analyzed to determine if they were still able to produce protein and if they had retained their DNA-binding properties. All mutated coding sequences were successfully translated *in vitro* by a reticulocyte lysate system (Fig 20).

Previous data from Mallo's lab had shown that Hox proteins were able to bind to a region in the promoter of the *Lbx1* gene. In order to evaluate DNA-binding ability of the mutated proteins, the same *Lbx1* DNA probe was used to conduct an electrophoretic mobility shift assay (EMSA) using the reticulocyte lysates. All proteins, including chimeric and mutant constructs, conserved the ability to bind to this *Lbx1* promoter region (Fig 21).

These results indicate that *in vitro*, the different phenotypes obtained were not a result of differential DNA-binding ability.

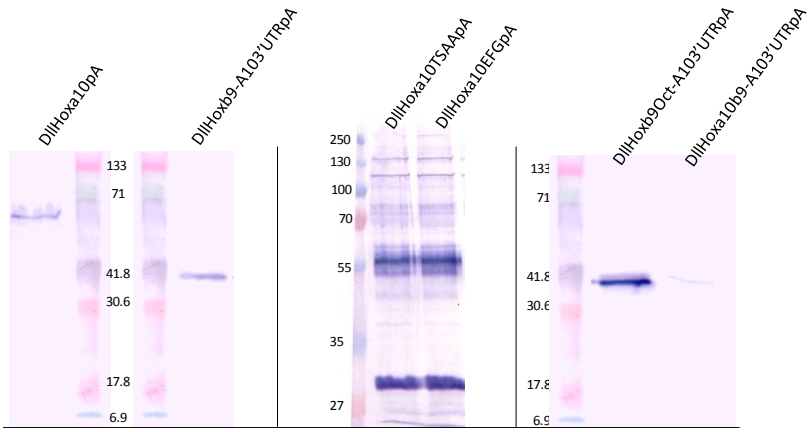


Figure 20 Western blot for proteins produced by a reticulocyte lysate system.
Both the unaltered proteins and mutated proteins are represented. Protein standard molecular weights are in KDa.

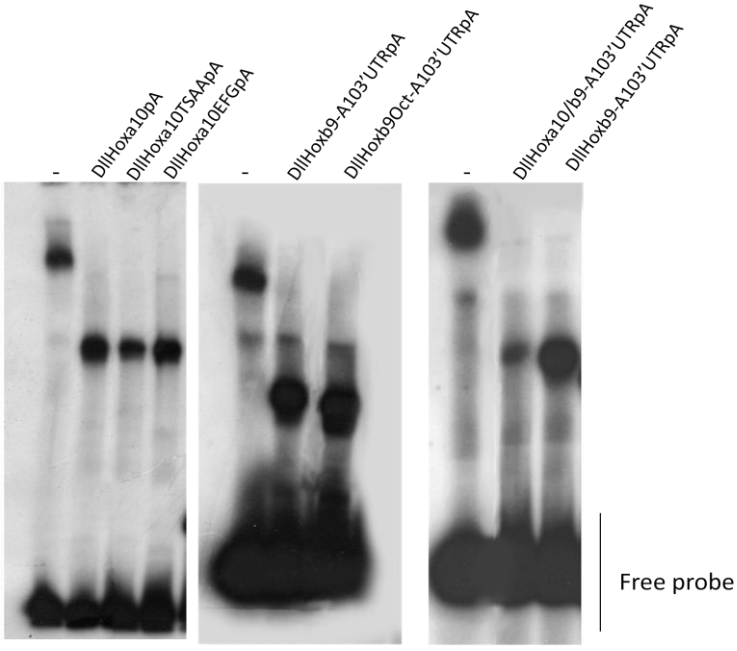


Figure 21 Electrophoretic mobility shift assay.
Both the unaltered and mutated proteins produced *in vitro* bind to an *Lbx1* DNA probe.

IV. Discussion

This work has addressed protein signatures that give specificity to the Hox group 10 proteins in their axial patterning function.

HD swapping experiments in mice have been performed in order to test if the HDs were functionally interchangeable. The swap of the Hoxa11 HD for that of the Hoxa13, Hoxa10 and Hoxa4 indicated that some functions are affected by the swap of a different homeodomain. However, the axial skeleton was not significantly affected in any of the mutant mice. Apparently, the HD function in the axial skeleton has been highly conserved, while its function in other tissues has diverged over time (Zhao *et al.*, 2001, Zhao *et al.*, 2002). Previous reports have showed that HD amino acids play an important role in Hox protein specificity (Chan *et al.*, 1993). However, the results obtained in these HD swapping experiments strongly suggest that the Hoxa10 HD is not likely to have a role in its axial patterning functions.

Some paralogs, including Hox group 10, have been reported to possess signature residues adjacent to the HD and very few conserved in the whole paralogous group within the HD. It is therefore likely that these peptide motifs outside the HD contribute to the functional specificity that does not seem to be provided by the HD alone (Sharkey et al., 1997). In fact, the removal of the conserved Hox10 octapeptide impaired almost completely Hox10's rib-repressing activity, showing that sequences outside the HD are necessary for Hox protein function. In addition, the insertion of this octapeptide in the Hoxb9 protein caused an abnormal axial skeleton phenotype, confirming the importance of this octapeptide. However, only mild axial skeleton phenotypes were obtained, suggesting that, in addition to the octapeptide, other residues within or outside the HD could be necessary for these proteins to carry out their patterning function. In particular, the other signature peptide motif of Hox group 10 proteins located just C-terminal to the HD could be necessary for Hoxb9 proteins to acquire a rib-repressing function. There is also the possibility that the octapeptide is not correctly positioned to perform the same function it does in Hoxa10. However, the Hoxb9 protein has a very similar peptide motif that differs only in three amino acids when compared to the Hoxa10 octapeptide. This makes a conformational issue unlikely to occur, although it does not fully exclude the possibility.

The presence of other residues lying N-terminally to the HD was also not sufficient to obtain a fully ribless phenotype. Instead they closely resemble the hypomorphic phenotype that results from low-copy *Hoxa10* overexpression in the PSM (Fig 11 b, e, h and Fig 19 b, d, f). For this reason, it is possible that the DllHoxa10/b9-A103'UTR chimeric contruct can, in fact, produce a more serious phenotype that has not been observed yet in the reduced number of transgenic embryos analyzed so far. In any case, the data provided so far indicates that C-terminal amino acids or the HD itself can also play a role in Hox10 patterning function. Experiments to test this hypothesis are underway. Two Hoxa10 constructs have been prepared: one of them codes for a Hoxa10 protein lacking the conserved peptide motif C-terminal to the HD and the other contains the Hoxb9 HD instead of its own. The overexpression of these constructs will hopefully clarify this matter.

There is strong evidence for the importance of the Hox10 octapeptide in Hox protein function. Pbx proteins have been shown to influence the DNA-binding ability of Hox proteins (Mann, 1997). The octapeptide analyzed here comprises the conserved tryptophan that has been shown to be essential for Hox-Pbx interaction (Chang *et al.*, 1996, Shen *et al.*, 1997). Since Pbx

proteins have been proposed to confer functional specificity to Hox proteins (Mann *et al.*, 1996), it is possible that the octapeptide is involved in providing specificity by interacting with Pbx. In order to test this idea, a construct with the tryptophan amino acid mutated to an alanine has already been generated, although it has not been microinjected yet. However, it has been previously documented that Pbx1 interactions might not be as relevant in AbdB-like Hox proteins (LaRonde-LeBlanc *et al.*, 2003). LaRonde-LeBlanc *et al.* described the structure of Hoxa9 and Pbx1 homeodomains bound to a DNA fragment. The Hoxa9 and Hoxa10 octapeptides were portrayed as hexapeptides with a three-amino acid linker to the HD to facilitate the comparison with the PBX-binding hexapeptide observed in Hox1-8 proteins. The conserved tryptophan residue preserved, as expected, its location in the Pbx pocket. Otherwise the divergent Abd-B-like hexapeptide was found to have a significantly different conformation compared with other documented structures. Although Pbx1 increased DNA-binding specificity of more anterior expressing Hox proteins such as Hoxb1, it did not do so for Hoxa9 and presumably other Abd-B related proteins such as Hoxa10 (LaRonde-LeBlanc *et al.*, 2003). It is therefore likely the mechanism responsible for providing Hoxa10 rib-repressing activity is not dependent on binding to Pbx1. Binding to other co-factors, post-translational modifications or recruitment of proteins that modulate chromatin structure are far more likely possibilities.

Protein phosphorylation is an essential mechanism for the regulation of many cellular functions like metabolism, proliferation, differentiation and apoptosis (Brinkworth et al., 2003). It has been reported that Hox genes can also have their activity modified as a result of phosphorylation (Jaffe et al., 1997). Therefore, it is possible that Hox10 proteins could have their rib-repressing activity regulated by phosphorylation, as well. The results obtained in this work seem to be consistent with this hypothesis, since mutations in the phosphorylation sites blocked the activity of Hoxa10 in inhibiting rib formation. However, phosphorylation has also been associated with ubiquitination, which suggests that protein degradation can also be part of the mechanism (Dimmeler *et al.*, 1999, Chang *et al.*, 1998). Although the protein was successfully produced in two cell types, a western blot should be conducted using transgenic embryonic tissue to confirm that the mutated protein is not being degraded *in vivo*. Surprisingly, C3H-10T1/2 cells show a significantly smaller band in the western blot, which could indicate the presence of a different splice variant. This unexpected result has not been further explored yet due to time constraints.

In order to determine if the Hoxa10 protein is being phosphorylated, a two-dimensional polyacrylamide gel electrophoresis (2D PAGE), that combines both isoelectric focusing and SDS-PAGE, will be conducted. This method allows the detection of specific qualitative and quantitative protein changes, such as post-translational modifications. The phosphorylation alters both the protein molecular weight and its isoelectric point. A difference in the 2D gel patterns obtained with the normal Hoxa10 protein and the Hoxa10 protein with both possible phosphorylation sites mutated could indicate that these residues, conserved throughout the whole paralogous group 10 proteins, are being phosphorylated.

All mutated and chimeric constructs were shown to be able to produce protein and bind to DNA. However, the Hoxa10/b9 construct showed a weak band in the EMSA which cannot be interpreted as a decrease in DNA-binding ability (Fig 21). It simply means that there was less protein synthesized *in vitro* and, consequently, less protein available to bind to the radioactive probe (Fig 20).

This study strongly suggests that sequences outside the HD have a critical role in Hox protein functional specificity. In this case, an octapeptide just upstream of the HD seems to be necessary for Hoxa10's rib-repressing activity. It is however insufficient to produce a full ribless phenotype on its own. It remains to be seen what other residues, within or outside the HD, are necessary for this transcription factor to achieve its functional specificity. The study through mutated and chimeric proteins is an important initial step to understand the mechanisms by which Hox genes operate on their downstream targets and ultimately establish differences along the AP axis.

References

Aoyama, H. and Asamoto, K. (1988). Determination of somite cells: independence of cell differentiation and morphogenesis. *Development* **104,** 15-28.

Bateson, W. (1984) Materials for the study of variation treated with especial regard to discontinuity in the origin of species Macmillan, London.

Beckers, J., Caron, A., Hrabe 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 *Mechanisms of development* **95,** 23-34.

Berthelsen, J., Zappavigna, V., Mavilio, F. and Blasi, F. (1998). Prep1, a novel functional partner of Pbx proteins. *The EMBO Journal* **17,** 1423-1433.

Bettenhausen, B., Hrabe de Angelis, M., Simon, D., Guenet, J.L. and Gossler, A. (1995). Transient and restricted expression during mouse embryogenesis of Dll1, a murine gene closely related to Drosophila Delta. *Development* **121,** 2407-2418.

Biggin, M.D. and McGinnis, W. (1997). Regulation of segmentation and segmental identity by Drosophila homeoproteins: the role of DNA binding in functional activity and specificity. *Development* **124,** 4425-4433.

Boncinelli, E., Simeone, A., La Volpe, A., Faiella, A., Fidanza, V., Campora, D. and Scotto, L. (1985). Human cDNA clones containing homeobox sequences. *Cold Spring Harb. Symp. quant. Biol* **50,** 301-306.

Brent, A.E. and Tabin, C.J. (2002). Developmental regulation of somite derivatives: muscle, cartilage and tendon. *Current Opinion in Genetics & Development* **12,** 548-557.

Brinkworth, R.I., Breinl, R.A. and Kobe, B. (2003). Structural basis and prediction of substrate specificity in protein serine/threonine kinases. *Proceedings of the National Academy of Sciences of the United States of America* **100,** 74-79.

Buckingham, M. (2001). Skeletal muscle formation in vertebrates. *Curr Opin Genet Dev* **11,** 440-448.

Burglin, T.R. (1997). Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals. *Nucl. Acids Res.* **25,** 4173-4180.

Burke, A.C. (2000). Hox genes and the global patterning of the somitic mesoderm. *Curr Top Dev Biol* **47,** 155-181.

Burke, A.C., Nelson, C.E., Morgan, B.A. and Tabin, C. (1995). Hox genes and the evolution of vertebrate axial morphology. *Development* **121,** 333-346.

Carapuço, M., Névoa, A., Bobola, N. and Mallo, M. (2005). Hox genes specify vertebral types in the presomitic mesoderm. *Genes & Development* **19,** 2116-2121.

Carrasco, A., McGinnis, W., Gehring, W. and De Robertis, E. (1984). Cloning of an X. laevis gene expressed during early embryogenesis coding for a peptide region homologous to Drosophila homeotic genes. *Cell* **37,** 409-414.

Chan, S.K. and Mann, R.S. (1993). The segment identity functions of Ultrabithorax are contained within its homeo domain and carboxy-terminal sequences. *Genes & Development* **7,** 796-811.

Chan, S.K. and Mann, R.S. (1996). A structural model for a homeotic protein-extradenticle-DNA complex accounts for the choice of HOX protein in the heterodimer. *Proceedings of the National Academy of Sciences of the United States of America* **93,** 5223-5228.

Chang, C.P., Brocchieri, L., Shen, W.F., Largman, C. and Cleary, M.L. (1996). Pbx modulation of Hox homeodomain amino-terminal arms establishes different DNA-binding specificities across the Hox locus. *Mol. Cell. Biol.* **16,** 1734-1745.

Chang, C.P., Shen, W.F., Rozenfeld, S., Lawrence, H.J., Largman, C. and Cleary, M.L. (1995). Pbx proteins display hexapeptide-dependent cooperative DNA binding with a subset of Hox proteins. *Genes & Development* **9,** 663-674.

Chang, W.-T., Thomason, P.A., Gross, J.D. and Newell, P.C. (1998). Evidence that the RdeA protein is a component of a multistep phosphorelay modulating rate of development in Dictyostelium. *EMBO J* **17,** 2809-2816.

Chauvet, S., Merabet, S., Bilder, D., Scott, M.P., Pradel, J. and Graba, Y. (2000). Distinct Hox protein sequences determine specificity in different tissues. *Proceedings of the National Academy of Sciences of the United States of America* **97,** 4064-4069.

Cordes, R., Schuster-Gossler, K., Serth, K. and Gossler, A. (2004). Specification of vertebral identity is coupled to Notch signalling and the segmentation clock. *Development* **131,** 1221-1233.

Dale, K. and Pourquié, O. (2000). A clock-work somite. *BioEssays* **22,** 72-83.

Deschamps, J., Van den Akker, E., Forlani, S., 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.

Deschamps, J. and van Nes, J. (2005). Developmental regulation of the Hox genes during axial morphogenesis in the mouse. *Development* **132,** 2931-2942.

Dimmeler, S., Breitschopf, K., Haendeler, J. and Zeiher, A.M. (1999). Dephosphorylation Targets Bcl-2 for Ubiquitin-dependent Degradation: A Link between the Apoptosome and the Proteasome Pathway. *J. Exp. Med.* **189,** 1815-1822.

Duboule, D. (1998). Vertebrate Hox gene regulation: clustering and/or colinearity? *Current Opinion in Genetics & Development* **8,** 514-518.

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,** 1497–1505.

Duboule, D. and Morata, G. (1994). Colinearity and functional hierarchy among genes of the homeotic complexes. *Trends in Genetics* **10,** 358-364.

Dubrulle, J. and Pourquie, O. (2004). Coupling segmentation to axis formation. *Development* **131,** 5783-5793.

Ekker, S.C., Young, K.E., Vonkessler, D.P. and Beachy, P.A. (1991). Optimal DNA-sequence recognition by the Ultrabithorax homeodomain of Drosophila. *Embo Journal* **10,** 1179-1186.

Favier, B. and Dolle, P. (1997). Developmental functions of mammalian Hox genes. *Mol. Hum. Reprod.* **3,** 115-131.

Fromental-Ramain, C., Warot, X., Lakkaraju, S., Favier, B., Haack, H., Birling, C*., et al.* (1996). Specific and redundant functions of the paralogous Hoxa-9 and Hoxd-9 genes in forelimb and axial skeleton patterning. *Development* **122,** 461-472.

Furukubo-Tokunaga, K., Flister, S. and Gehring, W.J. (1993). Functional specificity of the Antennapedia homeodomain. *Proceedings of the National Academy of Sciences of the United States of America* **90,** 6360-6364.

Gaunt, S. and Strachan, L. (1994). Forward spreading in the establishment of a vertebrate Hox expression boundary: The expression domain separates into anterior and posterior zones, and the spread occurs across implanted glass barriers. *Developmental Dynamics* **199,** 229-240.

Gaunt, S.J. (1988). Mouse homeobox gene transcripts occupy different but overlapping domains in embryonic germ layers and organs: a comparison of Hox-3.1 and Hox-1.5. *Development* **103,** 135-144.

Gehring, W., Müller, M., Affolter, M., Percival-Smith, A., Billeter, M., Qian, Y*., et al.* (1990). The structure of the homeodomain and its functional implications. *Trends in Genetics* **6,** 323-329.

Gehring, W.J. (1987). Homeo boxes in the study of development. *Science* **236,** 1245-1252.

Gilbert, S.F. (2006) Developmental biology Sunderland, MA, Sinauer Associates Inc.

Greer, J.M., Puetz, J., Thomas, K.R. and Capecchi, M.R. (2000). Maintenance of functional equivalence during paralogous Hox gene evolution. *Nature* **403,** 661-665.

Hart, C., Awgulewitsch, A., Fainsod, A., McGinnis, W. and Ruddle, F. (1985). Homeo box gene complex on mouse chromosome 11: Molecular cloning, expression in embryogenesis, and homology to a human homeo box locus. *Cell* **43,** 9-18.

Jaffe, L., Ryoo, H.D. and Mann, R.S. (1997). A role for phosphorylation by casein kinase II in modulating Antennapedia activity in Drosophila. *Genes & Development* **11,** 1327-1340.

Joshi, R., Passner, J.M., Rohs, R., Jain, R., Sosinsky, A., Crickmore, M.A.*, et al.* (2007). Functional specificity of a Hox protein mediated by the recognition of minor groove structure. *Cell* **131,** 530-543.

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.

Kissinger, C., Liu, B., Martin-Blanco, E., Kornberg, T. and Pabo, C. (1990). Crystal structure of an engrailed homeodomain-DNA complex at 2.8 A resolution: a framework for understanding homeodomain-DNA interactions. *Cell* **63,** 579-590.

Kmita, M. and Duboule, D. (2003). Organizing Axes in Time and Space; 25 Years of Colinear Tinkering. *Science* **301,** 331-333.

LaRonde-LeBlanc, N. and Wolberger, C. (2003). Structure of HoxA9 and Pbx1 bound to DNA: Hox hexapeptide and DNA recognition anterior to posterior. *Genes and Development* **17,** 2060-2072.

Lewis, E.B. (1978). Gene complex controlling segmentation in Drosophila. *Nature* **276,** 565-570.

Li, X. and McGinnis, W. (1999). Activity regulation of Hox proteins, a mechanism for altering functional specificity in development and evolution. *Proceedings of the National Academy of Sciences of the United States of America* **96,** 6802-6807.

Lin, L. and McGinnis, W. (1992). Mapping functional specificity in the Dfd and Ubx homeo domains. *Genes & Development* **6,** 1071-1081.

Lu, Y., Goldenberg, I., Bei, L., Andrejic, J. and Eklund, E.A. (2003). HoxA10 Represses Gene Transcription in Undifferentiated Myeloid Cells by Interaction with Histone Deacetylase 2. *Journal of Biological Chemistry* **278,** 47792-47802.

Mahmoudi, T. and Verrijzer, P. (2001). Chromatin silencing and activation by Polycomb and trithorax group proteins. *Oncogene* **20,** 3055-3066.

Mallo, M. (2007). And the segmentation clock keeps ticking. *BioEssays* **29,** 412-415.

Mallo, M., Vinagre, T. and Carapuço, M. (2008). The road to the vertebral formula. *Int J Dev Biol* **52**.

Mann, R. (1997). Why are Hox genes clustered? *BioEssays* **19,** 661-664.

Mann, R.S. and Affolter, M. (1998). Hox proteins meet more partners. *Current Opinion in Genetics & Development* **8,** 423-429.

Mann, R.S. and Chan, S.K. (1996). Extra specificity from *extradenticle*: the partnership between Hox and Pbx/Exd homeodomain proteins. *Trends in Genetics* **12,** 258-262.

McIntyre, D.C., Rakshit, S., Yallowitz, A.R., Loken, L., Jeannotte, L., Capecchi, M.R. and Wellik, D.M. (2007). Hox patterning of the vertebrate rib cage. *Development* **134,** 2981-2989.

Merabet, S., Hudry, B., Saadaoui, M. and Graba, Y. (2009). Classification of sequence signatures: a guide to Hox protein function. *BioEssays* **31,** 500-511.

Merabet, S., Kambris, Z., Capovilla, M., Berenger, H., Pradel, J. and Graba, Y. (2003). The hexapeptide and linker regions of the AbdA Hox protein regulate its activating and repressive functions. *Developmental Cell* **4,** 761-768.

Moens, C. and Selleri, L. (2006). Hox cofactors in vertebrate development *Developmental Biology* **291,** 193-206.

Monsoro-Burq, A.H. (2005). Sclerotome development and morphogenesis: when experimental embryology meets genetics. *Int J Dev Biol* **49,** 301-308.

Otting, G., Qian, Y., Billeter, M., Müller, M., Affolter, M., Gehring, W. and Wüthrich, K. (1990). Protein-DNA contacts in the structure of a homeodomain-DNA complex determined by nuclear magnetic resonance spectroscopy in solution. *Embo Journal* **9,** 3085-3092.

Palmeirim, I., Henrique, D., Ish-Horowicz, D. and Pourquié, O. (1997). Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. *Cell* **91,** 639-648.

Pearson, J., Lemons, D. and McGinnis, W. (2005). Modulating Hox gene functions during animal body patterning. *Nature reviews genetics* **6,** 893-904.

Pourquie, O. (2003). The Segmentation Clock: Converting Embryonic Time into Spatial Pattern. *Science* **301,** 328-330.

Prince, F., Katsuyama, T., Oshima, Y., Plaza, S., Resendez-Perez, D., Berry, M*., et al.* (2008). The YPWM motif links Antennapedia to the basal transcriptional machinery. *Development* **135,** 1669-1679.

Prince, V. (2002). The Hox Paradox: More Complex(es) Than Imagined *Developmental Biology* **249,** 1-15.

Rieckhof, G.E., Casares, F., Ryoo, H.D., Abu-Shaar, M. and Mann, R.S. (1997). Nuclear translocation of extradenticle requires homothorax, which encodes an extradenticle-related homeodomain protein. *Cell* **91,** 171-183.

Saleh, M., Rambaldi, I., Yang, X.-J. and Featherstone, M.S. (2000). Cell Signaling Switches HOX-PBX Complexes from Repressors to Activators of Transcription Mediated by Histone Deacetylases and Histone Acetyltransferases. *Mol. Cell. Biol.* **20,** 8623-8633.

Schier, A.F. and Gehring, W.J. (1993). Functional specificity of the homeodomain protein fushi tarazu: the role of DNA-binding specificity in vivo. *Proceedings of the National Academy of Sciences of the United States of America* **90,** 1450-1454.

Sharkey, M., Graba, Y. and Scott, M. (1997). Hox genes in evolution: protein surfaces and paralog groups *Trends in Genetics* **13,** 145-151.

Shen, W.-f., Krishnan, K., Lawrence, H.J. and Largman, C. (2001). The HOX Homeodomain Proteins Block CBP Histone Acetyltransferase Activity. *Mol. Cell. Biol.* **21,** 7509-7522.

Shen, W.-F., Rozenfeld, S., Lawrence, H.J. and Largman, C. (1997). The Abd-B-like Hox Homeodomain Proteins Can Be Subdivided by the Ability to Form Complexes with Pbx1a on a Novel DNA Target. *J. Biol. Chem.* **272,** 8198-8206.

Slack, J.M.W., Holland, P.W.H. and Graham, C.F. (1993). The zootype and the phylotypic stage. *Nature* **361,** 490-492.

Tam, P.P. and Beddington, R.S. (1987). The formation of mesodermal tissues in the mouse embryo during gastrulation and early organogenesis. *Development* **99,** 109-126.

Wellik, D.M. (2007). Hox patterning of the vertebrate axial skeleton. *Developmental Dynamics* **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,** 363-367.

Zeng, W., Andrew, D.J., Mathies, L.D., Horner, M.A. and Scott, M.P. (1993). Ectopic expression and function of the Antp and Scr homeotic genes: the N terminus of the homeodomain is critical to functional specificity. *Development* **118,** 339-352.

Zhao, Y. and Potter, S. (2001). Functional specificity of the Hoxa13 homeobox. *Development* **128,** 3197-3207.

Zhao, Y. and Potter, S. (2002). Functional Comparison of the Hoxa 4, Hoxa 10, and Hoxa 11 Homeoboxes. *Developmental Biology* **244,** 21-36.

APPENDIX I – BUFFERS, SOLUTIONS AND MEDIA

BUFFERS FOR MULTIPLE USES

1x TE	
EDTA	1mM
Tris-HCl	10mM

1xTAE	
EDTA (pH 8)	1mM
Acetic acid	20mM
Tris base	40mM

Gel loading buffer 6x	
Glycerol	30%
Bromophenol blue	0.25%

Ladders	
DNA ladder	Fermentas #SM0331
Protein ladder	Fermentas #SM1819
	Bio-Rad #161-0324

TO GENERATE COMPETENT CELLS

RF1	
RbCl	100 mM
MnCl2.4H2O	50 mM
Potassium Acetate	30 mM
CaCl2.2H20	10 mM
Glycerol	15 % (w/v)
	Adjust to pH 5.8 with 0.2M acetic acid
	Sterilized by filtration

RF2	
MOPS	10 mM
RbCl	10 mM
CaCl2.2H20	75 mM
Glycerol	15 % (w/v)
	Adjust to pH 6.8 with NaOH
	Sterilized by filtration

BACTERIAL GROWTH AND PLASMID PURIFICATION

Lysogeny Broth (LB) medium

Tryptone	1%
Yeast extract	0.5%
NaCl	1%

TENS

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

PCR

Enzymes

Pfu DNA polymerase	Fermentas (#EP0572)
Taq DNA polymerase	Fermentas (#EP0281)

10 x Buffers provided with the enzymes

KITS

DNA Kits

QIAEX II Gel Extraction kit	QIAGEN (#20051)
QIAprep spin miniprep kit	QIAGEN (#27104)
QIAquick gel extraction kit	QIAGEN (#28706)
QIAquick PCR Purification Kit	QIAGEN (#28106)
Plasmid DNA purification kit	NucleoBond (#740573)

in vitro protein synthesis

TNT® Sp6 reticulocyte lysate system	Promega (#L4601)

GENOTYPING

Laird's buffer

Tris-HCl, pH 8.5	100mM
EDTA	5mM
SDS	0.2%
NaCl	200mM

AppendixI

SKELETAL STAINING

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

Alizarin red solution	
Alizarin red S	50mg/L
KOH	2%

WESTERN

Lysis Buffer	
1 M NaCl	7.5 mL
10% NP-40	5 mL
20% SDS	0.25 mL
1M Tris-HCl pH 8.0	2.5 mL

10x Tris-Glycine	
Tris base	25mM
Glycine	192mM

Running buffer	
Tris-Glycine	1X
SDS	0,1%

Transfer buffer	
Tris-Glycine	1X
Methanol	20%

2X SDS-containing loading buffer	
Tris-HCl pH 6.8	125mM
SDS	4%
Glycerol	20%
Bromophenol blue	0.006%
beta-mercaptoethanol	1.8%

10-12% SDS polyacrylamide resolving gel	
30%acrylamide/bisacrylamide	10%
1.5 M Tris pH 8.8	390mM
20% SDS	0.05%
10% ammonium persulfate	0.1%
TEMED	0.04%

5% SDS polyacrylamide stacking gel	
30%acrylamide/bisacrylamide	5%
1M Tris pH 6.8	125mM
20% SDS	0.05%
10% ammonium persulfate	0.1%
TEMED	0.04%

Blocking Buffer	
BSA	3%
10% Tween 20	0.1%
PBS	to final vol

1x PBS	
NaCl	137 mM
KCl	2.7 mM
Na2HPO4	10 mM
KH2PO4	2 mM
	Adjust pH to 7.4 with HCl

CELL CULTURE

Growth media	
DMEM	Sigma (#D5796)
Fetal Calf serum	Sigma (#7524)
Penicillin and Streptomycin	Sigma (#P0781)
L-glutamine	Sigma (#G7513)

Trypsinization	
Trypsin with EDTA	Sigma (#T3924)
Dulbecco's PBS	Sigma (#D1408)

Transfection	
Lipofectamin™ 2000	Invitrogen (#11668-019)

EMSA

Probe labelling	
T4 Polynucleotide Kinase	Promega (#M4101)
T4 PNK Buffer	Promega (#C1313)

5XTBE	
Tris Base	445 mM
boric acid	445 mM

EDTA (pH 8.0)	10 mM

5% non-denaturing polyacrylamide gel

30%acrylamide/bisacrylamide	5%
TBE	0.5x
10% ammonium persulfate	0.1%
TEMED	0.1%

APPENDIX II – CODING SEQUENCES OF TEMPLATES AND MUTATED CONSTRUCTS

AppendixII

All Hoxa10 sequences are uppercase and Hoxb9 sequences are in lowercase. The octapepide is underlined. Amino acid point mutations are in bold.

Hoxa10

ATGTCATGCTCGGAGAGCCCTGCCGCGAACTCCTTTTTGGTCGACTCGCTCATCAGCTCAGGCAGAGGC
GAGGCTGGTGGTGGTGGCGGTAGCGCGGGGGGCGGTGGAGGTGGCTACTACGCCCACGGTGGGGTC
TACCTGCCGCCTGCCAGCGACCTGCCCTACGGGCTGCAAAGCTGCGGGCTCTTCCCCGCGCTGGGCAG
CAAGCGTAATGAAGCGCCGTCGCCCGGAGGCGGTGGCGGTGGTGGCAGCGGGGGCCTGGGTCCTGG
GACGCATGGCTACGCGCCCGCGCCCCTAGACCTGTGGCTGGACGCGCCCCGCTCCTGCCGGATGGAGC
CGCCCGACGGGCCGCCGCCACCGCAGCCACAACCCCAGCAGCAGCAGCAGCAGCCGCCGCCGCCCCC
GCCGCAGCCACCTCAACCCCAGCCACAGGCCACTTCGTGTTCTTTTGCGCAGAACATCAAAGAAGAGAG
CTCCTACTGCCTCTACGATGCTGCGGACAAATGCCCCAAGGGCTCGGCCGCCGCTGATCTGGCCCCTTT
CCCGCGGGGCCCGCCGCCCGACGGCTGCGCCCTGGGCGCCTCCAGCGGAGTGCCAGTACCCGGCTACT
TCCGCCTGTCGCAGGCCTACGGCACGGCCAAGGGCTTCGGCAGTGGCGGCGGCGGCACGCAGCAGCT
CGCTAGTCCCTTTCCTGCGCAGCCCCCGGGGCGCGGTTTCGACCCGCCGCCCGCACTGGCCTCTGGCTC
GACCGAGGCAGCCGGGAAGGAGCGAGTCCTAGACTCCACGCCACCACCCACTCTGGTTTGCACCGGTG
GCGGCGGCTCGCAGGGCGACGAGGAGGCACACGCGTCATCCTCGGCGGCTGAGGAGCTGTCTCCAGC
CCCTTCAGAAAACAGTAAAGCTTCGCCGGAGAAGGACTCCCTGGGCAGTTCCAAAGGCGAAAATGCAG
CCAACTGGCTCACAGCAAAGAGCGGCCGGAAGAAACGCTGCCCTTACACGAAGCACCAGACGCTGGA
GCTGGAGAAGGAGTTTCTATTCAACATGTACCTTACTCGAGAGCGGCGCCTAGAGATCAGCCGTAGCG
TCCACCTCACGGACAGACAAGTGAAAATCTGGTTTCAGAATCGCAGGATGAAACTGAAGAAAATGAAC
CGAGAAAACCGAATCCGGGAGCTCACAGCCAACTTTAATTTTTCC

Hoxa10EFG

ATGTCATGCTCGGAGAGCCCTGCCGCGAACTCCTTTTTGGTCGACTCGCTCATCAGCTCAGGCAGAGGC
GAGGCTGGTGGTGGTGGCGGTAGCGCGGGGGGCGGTGGAGGTGGCTACTACGCCCACGGTGGGGTC
TACCTGCCGCCTGCCAGCGACCTGCCCTACGGGCTGCAAAGCTGCGGGCTCTTCCCCGCGCTGGGCAG
CAAGCGTAATGAAGCGCCGTCGCCCGGAGGCGGTGGCGGTGGTGGCAGCGGGGGCCTGGGTCCTGG
GACGCATGGCTACGCGCCCGCGCCCCTAGACCTGTGGCTGGACGCGCCCCGCTCCTGCCGGATGGAGC
CGCCCGACGGGCCGCCGCCACCGCAGCCACAACCCCAGCAGCAGCAGCAGCAGCCGCCGCCGCCCCC
GCCGCAGCCACCTCAACCCCAGCCACAGGCCACTTCGTGTTCTTTTGCGCAGAACATCAAAGAAGAGAG
CTCCTACTGCCTCTACGATGCTGCGGACAAATGCCCCAAGGGCTCGGCCGCCGCTGATCTGGCCCCTTT
CCCGCGGGGCCCGCCGCCCGACGGCTGCGCCCTGGGCGCCTCCAGCGGAGTGCCAGTACCCGGCTACT
TCCGCCTGTCGCAGGCCTACGGCACGGCCAAGGGCTTCGGCAGTGGCGGCGGCGGCACGCAGCAGCT
CGCTAGTCCCTTTCCTGCGCAGCCCCCGGGGCGCGGTTTCGACCCGCCGCCCGCACTGGCCTCTGGCTC
GACCGAGGCAGCCGGGAAGGAGCGAGTCCTAGACTCCACGCCACCACCCACTCTGGTTTGCACCGGTG
GCGGCGGCTCGCAGGGCGACGAGGAGGCACACGCGTCATCCTCGGCGGCTGAGGAGCTGTCTCCAGC
CCCTTCAGAAAACAGTAAAGCTTCGCCGGAGAAGGACTCCCTGGGCAGTTCCAAAGGCGAAAATGCAG
CCCGGAAGAAACGCTGCCCTTACACGAAGCACCAGACGCTGGAGCTGGAGAAGGAGTTTCTATTCAAC
ATGTACCTTACTCGAGAGCGGCGCCTAGAGATCAGCCGTAGCGTCCACCTCACGGACAGACAAGTGAA
AATCTGGTTTCAGAATCGCAGGATGAAACTGAAGAAAATGAACCGAGAAAACCGAATCCGGGAGCTCA
CAGCCAACTTTAATTTTTCC

Hoxa10TSAA

ATGTCATGCTCGGAGAGCCCTGCCGCGAACTCCTTTTTGGTCGACTCGCTCATCAGCTCAGGCAGAGGC
GAGGCTGGTGGTGGTGGCGGTAGCGCGGGGGGCGGTGGAGGTGGCTACTACGCCCACGGTGGGGTC
TACCTGCCGCCTGCCAGCGACCTGCCCTACGGGCTGCAAAGCTGCGGGCTCTTCCCCGCGCTGGGCAG
CAAGCGTAATGAAGCGCCGTCGCCCGGAGGCGGTGGCGGTGGTGGCAGCGGGGGCCTGGGTCCTGG
GACGCATGGCTACGCGCCCGCGCCCCTAGACCTGTGGCTGGACGCGCCCCGCTCCTGCCGGATGGAGC
CGCCCGACGGGCCGCCGCCACCGCAGCCACAACCCCAGCAGCAGCAGCAGCAGCCGCCGCCGCCCCC
GCCGCAGCCACCTCAACCCCAGCCACAGGCCACTTCGTGTTCTTTTGCGCAGAACATCAAAGAAGAGAG
CTCCTACTGCCTCTACGATGCTGCGGACAAATGCCCCAAGGGCTCGGCCGCCGCTGATCTGGCCCCTTT
CCCGCGGGGCCCGCCGCCCGACGGCTGCGCCCTGGGCGCCTCCAGCGGAGTGCCAGTACCCGGCTACT
TCCGCCTGTCGCAGGCCTACGGCACGGCCAAGGGCTTCGGCAGTGGCGGCGGCGGCACGCAGCAGCT
CGCTAGTCCCTTTCCTGCGCAGCCCCCGGGGCGCGGTTTCGACCCGCCGCCCGCACTGGCCTCTGGCTC
GACCGAGGCAGCCGGGAAGGAGCGAGTCCTAGACTCCACGCCACCACCCACTCTGGTTTGCACCGGTG
GCGGCGGCTCGCAGGGCGACGAGGAGGCACACGCGTCATCCTCGGCGGCTGAGGAGCTGTCTCCAGC
CCCTTCAGAAAACAGTAAAGCTTCGCCGGAGAAGGACTCCCTGGGCAGTTCCAAAGGCGAAAATGCAG
CCAACTGGCTC**GCA**GCAAAG**GCC**GGCCGGAAGAAACGCTGCCCTTACACGAAGCACCAGACGCTGGA
GCTGGAGAAGGAGTTTCTATTCAACATGTACCTTACTCGAGAGCGGCGCCTAGAGATCAGCCGTAGCG
TCCACCTCACGGACAGACAAGTGAAAATCTGGTTTCAGAATCGCAGGATGAAACTGAAGAAAATGAAC
CGAGAAAACCGAATCCGGGAGCTCACAGCCAACTTTAATTTTTCC

Hoxb9

atgtccatttctgggacgcttagcagctattatgtcgactcgatcataagtcacgagagtgaggacgcgcctccagccaagtttccttctg
gccagtacgcgagctcgcggcagccgggccacgcggagcacctggagttcccctcgtgcagcttccagcccaaagcgccggtgttcgg
cgcctcctgggcgccgctgagcccgcacgcgtccgggagcctgccgtccgtctaccacccttacatccagccccagggcgtcccgccgg
ccgagagcaggtacctccgcacctggctggagccggcgccgcgcggcgaagcggccccggggcagggccaggcggcggtgaaggcg
gagccgctgctgggcgcgcctggggagctgctcaaacagggcacgcccgagtacagtttggaaacttcggcgggcagggaggccgtg
ctgtctaatcaaagacccggctacggggacaataaaatttgcgaaggaagcgaggacaaagagaggccggatcaaaccaacccctcc
gccaactggctgcacgctcgctcttcccggaaaaagcgctgtccctacaccaaataccagacgctggagctagagaaggagtttctgtt
caatatgtacctcaccagggaccgtaggcacgaagtggccagactcctcaatctgagtgagagacaagtcaaaatctggtttcagaac
cggcggatgaaaatgaagaaaatgaataaggagcagggcaaagag

Hoxb9Oct

atgtccatttctgggacgcttagcagctattatgtcgactcgatcataagtcacgagagtgaggacgcgcctccagccaagtttccttctg
gccagtacgcgagctcgcggcagccgggccacgcggagcacctggagttcccctcgtgcagcttccagcccaaagcgccggtgttcgg
cgcctcctgggcgccgctgagcccgcacgcgtccgggagcctgccgtccgtctaccacccttacatccagccccagggcgtcccgccgg
ccgagagcaggtacctccgcacctggctggagccggcgccgcgcggcgaagcggccccggggcagggccaggcggcggtgaaggcg
gagccgctgctgggcgcgcctggggagctgctcaaacagggcacgcccgagtacagtttggaaacttcggcgggcagggaggccgtg
ctgtctaatcaaagacccggctacggggacaataaaatttgcgaaggaagcgaggacaaagagaggccggatcaaaccaacccctcc
gccAACTGGCTCACAGCAAAGAGCGGCcggaaaaagcgctgtccctacaccaaataccagacgctggagctagagaagg

agtttctgttcaatatgtacctcaccagggaccgtaggcacgaagtggccagactcctcaatctgagtgagagacaagtcaaaatctgg
tttcagaaccggcggatgaaaatgaagaaaatgaataaggagcagggcaaagag

Hoxa10b9

ATGCTCGGAGAGCCCTGCCGCGAACTCCTTTTTGGTCGACTCGCTCATCAGCTCAGGCAGAGGCGAGGCT
GGTGGTGGTGGCGGTAGCGCGGGGGGCGGTGGAGGTGGCTACTACGCCCACGGTGGGGTCTACCTGCCGC
CTGCCAGCGACCTGCCCTACGGGCTGCAAAGCTGCGGGCTCTTCCCCGCGCTGGGCAGCAAGCGTAATGA
AGCGCCGTCGCCCGGAGGCGGTGGCGGTGGTGGCAGCGGGGGCCTGGGTCCTGGGACGCATGGCTACGCG
CCCGCGCCCCTAGACCTGTGGCTGGACGCGCCCCGCTCCTGCCGGATGGAGCCGCCCGACGGGCCGCCGC
CACCGCAGCCACAACCCCAGCAGCAGCAGCAGCCGCCGCCGCCCCCGCCGCAGCCACCTCAACCCCA
GCCACAGGCCACTTCGTGTTCTTTTGCGCAGAACATCAAAGAAGAGAGCTCCTACTGCCTCTACGATGCT
GCGGACAAATGCCCCAAGGGCTCGGCCGCCGCTGATCTGGCCCCTTTCCCGCGGGGCCCGCCGCCCGACG
GCTGCGCCCTGGGCGCCTCCAGCGGAGTGCCAGTACCCGGCTACTTCCGCCTGTCG**CAGGCCTACGGCAC**
GGCCAAGGGCTTCGGCAGTGGCGGCGGCGGCACGCAGCAGCTCGCTAGTCCCTTTCCTGCGCAGCCCCCG
GGGCGCGGTTTCGACCCGCCGCCCGCACTGGCCTCTGGCTCGACCGAGGCAGCCGGGAAGGAGCGAGTCC
TAGACTCCACGCCACCACCCACTCTGGTTTGCACCGGTGGCGGCGGCTCGCAGGGCGACGAGGAGGCACA
CGCGTCATCCTCGGCGGCTGAGGAGCTGTCTCCAGCCCCTTCAGAAAACAGTAAAGCTTCGCCGGAGAAG
GACTCCCTGGGCAGTTCCAAAGGCGAAAATGCAGC**CAACTGGCTCACAGCAAAGAGCGGCCGG**cggaaaa
agcgctgtccctacaccaaataccagacgctggagctagagaaggagtttctgttcaatatgtacctcac
cagggaccgtaggcacgaagtggccagactcctcaatctgagtgagagacaagtcaaaatctggtttcag
aaccggcggatgaaaatgaagaaaatgaataaggagcagggcaaagag