Control of *Hox* gene regulation and function during anteroposterior patterning in *Xenopus laevis*

Controle van *Hox* gen regulatie en functie tijdens anteroposteriore patroonvorming in *Xenopus laevis*

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht,
op gezag van de Rector Magnificus, Prof. Dr. W. H. Gispen,
ingevolge het besluit van het College voor Promoties
in het openbaar te verdedigen op
woensdag 9 april 2003 des middags te 2:30 uur

door

Paul Marius Josephus In der Rieden

geboren op 20 mei 1974, te Emmen

Promotor: Prof. Dr. A. J. Durston

Hubrecht Laboratorium, Nederlands Instituut voor Ontwikkelingsbiologie, Uppsalalaan 8, 3584 CT, Utrecht, The Netherlands.

ISBN: 90-6734-206-8

This thesis presents investigations performed in the Hubrecht Laboratorium, Nederlands Instituut voor Ontwikkelingsbiologie, within the Graduate School for Developmental Biology and supported by a NWO-grant and EU "HPRN" and "Quality of life" programs.

Contents:

Chapter 1

Introduction

Paul M. J. In der Rieden

Hubrecht Laboratorium, Nederlands Instituut voor Ontwikkelingsbiologie,
Uppsalalaan 8, 3584 CT, Utrecht, The Netherlands.

Hox genes

The *Hox* genes form a subset of the homeobox containing genes. The homeobox encodes a DNA binding motif, called the homeodomain. In most animal species the *Hox* genes are organised in one or more clusters. The number of genes present in a cluster varies between animal species; eight in insects, ten in myriapods (Hughes and Kaufman, 2002), twelve in Amphioxus (Garcia-Fernàndez and Holland, 1994), and up to thirteen in vertebrates (Acampora *et al.*, 1989; Graham *et al.*, 1989; Kim *et al.*, 2000). The number of clusters in each species also varies. Arthropods (Hughes and Kaufman, 2002) and amphioxus (Ruvkun and Hobert, 1998) posses one cluster, while vertebrates mostly contain four clusters (Acampora *et al.*, 1989, Graham *et al.*, 1989; Kim *et al.*, 2000), named HoxA through HoxD. The Hox clusters are thought to have arisen by tandem duplication of a single gene, followed, in vertebrates, by duplication of the cluster itself (Schughart *et al.*, 1989; Ruddle *et al.*, 1994; Bailey *et al.*, 1997). As a consequence, *Hox* genes occupying the same relative position along the 5' to 3' chromosomal coordinate, named paralogous genes, share more similarities in sequence, expression pattern and function than do adjacent *Hox* genes on the same chromosome (Condie *et al.*, 1994; Davis *et al.*, 1995; Horan *et al.*, 1995; Zákány *et al.*, 1996; Gavalas *et al.*, 1998; Struder *et al.*, 1998; Rossel and Capecchi, 1999; Greer *et al.*, 2000). One of the conserved features of the *Hox* genes, displayed generally in bilaterian embryos, is that of spatial colinearity of *Hox* expression along the AP axis and other embryonic axes (Gaunt, 1988; Duboule and Dollé, 1989; Graham *et al.*, 1989; Gaunt, 1991). This means that paralogs located in the most 3’ position of a cluster are expressed at a more anterior position than genes located at more 5’ clustral locations. Vertebrate embryos also display a second form of colinearity, temporal colinearity (Izpisua-Belmonte *et al.*, 1991; Duboule and Morata, 1994; Gaunt and Strachan, 1996; Deschamps *et al.*, 1999), whereby genes located more 3’ in the cluster are expressed earlier than genes located more 5’ in the cluster. As in vertebrates both colinearities are inseparable, they are referred to as spatiotemporal colinearity.
It has been firmly established that *Hox* expression boundaries along the AP and other embryonic axes are correlated with structural identities; misexpression of *Hox* genes can lead to patterning defects (reviewed by McGinnis and Krumlauf, 1992; Deschamps *et al.*, 1999; Gaunt, 2000). A textbook example of such an event is the *Antennapedia* mutation in *Drosophila*, whereby an antenna is transformed into a leg by ectopic expression of the *Antennapedia* gene in the eye-antennal disc (Jorgensen and

Garber, 1987; Schneuwly *et al.*, 1987). These homeotic transformations have also been observed as a feature of vertebrate *Hox* genes; ectopic expression in more anterior domains than the endogenous expression generally leads to posteriorisation (Kessel and Gruss, 1990; Kessel and Gruss 1991; Lufkin *et al.*, 1992). Generating correct *Hox* expression patterns is thus clearly essential for correct AP axis patterning. However, many questions about the regulation of *Hox* gene expression remain. In addition, the way in which the Hox transcription factors confer positional identity with high target specificity is still mysterious.

Aim and content of this thesis

The aim of investigations presented in this thesis has been to gain more insight into the processes controlling expression and function of *Hox* genes during anteroposterior patterning. This thesis can broadly be subdivided in three sections. The first section, chapters 2 and 3, deals with upstream regulators of *Hox* gene expression. The second section, chapters 4, 5, and 6, focuses on cofactors of Hox proteins, which are necessary for providing specificity and mediation of Hox transcriptional activation. Finally, in the third section, chapter 7, a downstream target of a Hox gene is described. Because the introduction of the chapter is complete it will not be dealt with in this introduction.

Factors upstream of *Hox* gene expression

Recent work in our research group has shown that a temporally colinear expression sequence of *Hox* genes is already present in the marginal zone mesoderm of *Xenopus* gastrulae (Wacker *et al.*, submitted; this thesis). Interactions between the ventrolateral mesoderm and the Spemann organiser bring the *Hox* cascade to a halt. This results in a pattern of mesodermal *Hox* expression whereby tissue that interacts with the organiser earlier during involution, continues to expresses the most anterior *Hox* genes, and tissue interacting with the organiser later expresses more posterior *Hox* genes. The authors show that a graft of organiser tissue, placed in an embryo ventralised by UV-irradiation, can restore a larger part of the AP axis if the graft is implanted earlier during gastrulation. If an identical graft was implanted at the end of gastrulation, i.e. in embryos expressing more posterior *Hox* genes, severe anterior truncations could be observed, the severity depending on the stage of implantation. This led to the conclusion that the progression of the *Hox* cascade in mesoderm is inhibited, at various points by interaction with

the organiser in order to pattern an AP axis. Most, if not all, previous work concerning *Hox* colinearity has been focused on colinear *Hox* gene expression in the neurectoderm. It is therefore of great interest to investigate which factors are upstream regulators of *Hox* gene expression in marginal zone mesoderm, we set out to identify factors involved in the regulation of *Hox* expression during these stages.

A factor upstream of *Hox* expression in marginal zone mesoderm

We argued that factors known to be involved in early mesodermal patterning are potential regulators of initiation and establishment of *Hox* expression. Taking into account that *Hox* genes, when overexpressed, have a posteriorising effect on embryos, we considered factors already known to be expressed in mesoderm, which are known to have a comparable effect on embryos. A factor capable of ventralising mesoderm and posteriorising neurectoderm is *Xwnt8* (Christian and Moon, 1993; Fredieu *et al.*, 1997; Erter *et al.*, 2001; Kiecker and Niehrs, 2001). We investigated whether *Xwnt8* could regulate *Hox* expression in marginal zone mesoderm, using gain- and loss-of-function strategies. In chapter 2, we present data to show that *Xwnt8* is directly upstream of *Hoxd1* in marginal zone mesoderm. This is the first example of an initiator of expression of a 3' *Hox* gene in a vertebrate. Interestingly, it employs a conserved signalling cascade found in a range of species varying from *C. elegans* to vertebrates.

A factor upstream of *Hox* gene expression in neurectoderm

An upstream regulator of *Hox* gene expression in the neurectoderm of vertebrates is retinoic acid (RA) (and or its derivatives) (Hofmann and Eichele, 1994; Durston *et al.*, 1998; Gavalas and Krumlauf, 2000). Retinoids can act via the nuclear receptors of the RAR and RXR family. These receptors form heterodimers and bind Retinoic Acid Response Elements (RAREs) in the promoters of target genes. In the absence of ligand they can act as transcriptional repressors (Cohen *et al.*, 2001); and addition of a suitable ligand turns them into activators. RAREs have been found in the regulatory sequences of a number of labial- and deformed group *Hox* genes. We took advantage of the availability of the genomic sequences of all four Hox clusters of mouse and human, to search for conserved RAREs in the Hox clusters, and the results are reported in chapter 3.

Hox function: interaction with cofactors

Hox proteins can achieve high target specificity *in vivo*, while *in vitro* DNA binding studies reveal low binding specificity and affinity (Hayashi and Scott, 1990; Laughon, 1991; Knoepfler *et al.*, 1996; Di Rocco *et al.*, 1997). Different target specificity of individual Hox proteins could be achieved by sequence variation in the DNA-contacting helix of the homeodomain. Strikingly, this helix, containing the amino acid sequence K-I-W-F-Q-N-R-R-M-R, is almost invariant among different Hox proteins from a wide range of species including *C. elegans* and human (Bürglin, 1994). However, It has been shown that target specificity is partly conferred by the N-terminal arm of the homeodomain when the Hox protein binds DNA as a partner in a heterodimer, formed with PBC-class cofactors (Chang *et al.*, 1996; Di Rocco *et al.*, 1997; Phelan and Featherstone, 1997). Promoter analysis and *in vitro* binding studies have revealed that Hox and PBC family proteins can cooperatively bind DNA (van Dijk and Murre, 1994; van Dijk *et al.*, 1995), on bipartite consensus sequences of the type TGATNNATNN (Chan and Mann, 1996; Chang *et al.*, 1996; Knoepfler *et al.*, 1999; Piper *et al.*, 1999). In addition, *in vivo* studies report that these bipartite binding sites are essential for Hox mediated activation of promoters that contain them (Pöpperl *et al.*, 1995; Chan *et al.*, 1997; Di Rocco *et al.*, 1997; Grieder *et al.*, 1997; Maconochie *et al.*, 1997; Ryoo and Mann, 1999; Ferretti *et al.*, 2000). The PBC family, encompassing the vertebrate Pbx proteins, *Drosophila* Exd, and *C. elegans* ceh-20, is a subfamily of the TALE class of homeodomain proteins, which contain an atypical, three amino acid extended homeodomain (reviewed by Bürglin, 1997). X-ray diffraction crystallography (Passner *et al.*, 1999; Piper *et al.*, 1999) and NMR studies (Jabet *et al.*, 1999) have shown that the Hox and PBC proteins bind target sequences as heterodimers. The PBC class member of the heterodimer binds the sequence GTATNN, while on the other strand of the binding element, the sequence NNATNN is recognised by the Hox partner (reviewed by Mann and Chan, 1996). These bipartite sequences have been found in the regulatory sequences of *Hox* genes themselves and have been shown to be essential for specific autoregulation by labial and deformed type *Hox* genes. Labial type Hox proteins preferably bind with PBC members to the sequence gtatGGatGG (Chan and Mann, 1996; Maconochie *et al.*, 1997; Mann and Affolter 1998; Piper *et al.*, 1999; Ferretti *et al.*, 2000), while deformed type Hox proteins/PBC heterodimers prefer the sequence gtatTAatGG (Gould *et al.*, 1997; Chan *et al.*, 1997; Mann and Affolter, 1998). This shows that the interaction of Hox proteins with PBC family cofactors leads to different

target specificity depending on the Hox partner of the heterodimer. The interaction between the PBC and Hox partners in DNA-bound heterodimers has been studied (Knoepfler and Kamps, 1995; Phelan *et al.*, 1995; Jabet *et al.*, 1999; Passner *et al.*, 1999; Piper *et al.*, 1999). The protein-protein interaction between Hox and PBC proteins consists of the binding of a conserved hexapeptide sequence, found in most Hox proteins and located N-terminal and proximal to the homeodomain, to a pocket formed by PBC family members. This pocket is composed of the three amino acid loop extension of PBC homeodomains, residues in helix 3 of the homeodomain, and a residue in the C-terminal helix of PBC homeodomains (Piper *et al.*, 1999). The hexapeptide, defined as a tryptophan residue in a hydrophobic context, flanked by arginine and lysine residues at +2 to +5 positions to the tryptophan (Knoepfler *et al.*, 1999), is found in all Hox proteins, with the exception of Hox 11 though 13 group members. The hexapeptide is not an exclusive hallmark of Hox proteins as it has also been found in the Engrailed homeodomain proteins (Peltenburg and Murre, 1996), and the myogenic bHLH transcription factors (Knoepfler *et al.*, 1999). The hexapeptide found in Hox group 1 through 8 proteins has the consensus: hydrophobic-Y/F-P-W-M-K/R (Piper *et al.*, 1999), while in Hox group 9 and 10 proteins a A-N-W-L/I-H/T-A consensus is found (Shen *et al.*, 1997). Although it is known that the hexapeptide is needed for a Hox protein to be able to interact with PBC members, and that this interaction leads to binding specificity, the fashion in which this specificity is achieved is unknown. Can target specificity be explained by variations in the hexapeptides of, for instance, labial (T-F-D-W-M-K) and deformed (V-Y-P-W-M-R) group members? Piper and co-workers (1999) have reported that interactions mediated by the first three residues of the labial hexapeptide could be mediated by the residues found in those positions in the deformed hexapeptide. These authors also report that the conservation of a basic residue at the sixth position cannot be explained by the crystal structure since it is disordered in the labial type Hox/Pbx/DNA complex they have studied. This suggests that the amino acid sequences of the two hexapeptides are functionally similar. In addition, it has been stated that naturally occurring variations among hexapeptides do not explain the difference in complex stabilities of various Hox proteins in Hox/Pbx/DNA combinations (Shen *et al.*, 1996). Together these results suggest that the individual sequences of hexapeptides themselves do not account for the specificity. To investigate whether additional paralog specific sequences might account for *in vivo* target specificity, we compared sequences of labial through Abd-A group Hox proteins. The discovery of interspecies sequence conservation among paralog group members in the sequences flanking the

hexapeptide is reported in chapter 4. Since sequence conservation between Hox proteins outside the homeodomain and hexapeptide is low, the conserved hexapeptide-flanking sequences are likely to be functional in defining paralog specificity. We propose that the interaction between Hox and PBC members is fine-tuned by the conserved hexapeptide-flanking regions, leading to different target specificities for different Hox/PBC combinations.
The activity of the paralog groups 1 through 10 Hox proteins is not only influenced at the level of Hox-PBC-DNA interactions. Nuclear localisation of PBC family members is controlled by competing nuclear import and export signals (Rieckhof *et al.*, 1997; Abu-Saar *et al.*, 1999). Meis proteins, also members of the TALE-class of homeodomain proteins, are the vertebrate homologs of *Drosophila* Homothorax and *C. elegans* ceh-25 (Bürglin, 1997). Interaction of Meis family members with PBC members shields the nuclear export signal of PBC proteins, resulting in a net influx into the nucleus, modifying the activity of Hox proteins present (Ryoo *et al.*, 1999; Ryoo and Mann, 1999; Jaw *et al.*, 2000). In vertebrate embryos, Meis proteins are essential for hindbrain patterning and they can posteriorise the neurectoderm when ectopically expressed (Salzberg *et al.*, 1999; Dibner *et al.*, 2001; Vlachakis *et al.*, 2001; Waskiewicz *et al.*, 2001; Cheo *et al.*, 2002). In zebrafish hindbrain development, a synergistic relation between Hoxb1, Pbx4, and Meis3 has been shown, and was argued to directly induce the expression of *Hoxb1* (Vlachakis *et al.*, 2001). Since recent discoveries have shown that *Hox* genes are expressed in a colinear sequence in marginal zone mesoderm (Wacker *et al.*, submitted; this thesis), we wished to investigate whether a *Xenopus* Meis homolog, *XMeis3*, cooperates with Hox function during gastrula stages. In chapter 5, we report that *XMeis3* is necessary for mesodermal and ectodermal *Hox* expression, and the progression of gastrulation.

Hexapeptide revisited

Several homeobox genes with high sequence similarities to *Hox* genes have been found not to be associated with the Hox clusters (Bürglin, 1994). In *Amphioxus*, three of these genes, *Gsx*, *Xlox*, and *Cdx*, are clustered (Brooke *et al.*, 1998). It has been proposed that this cluster, named the ParaHox cluster, is the evolutionary sister of the Hox cluster and that they both arose from a ProtoHox cluster. In human and mouse, orthologs of all three Amphioxus *ParaHox* genes, are localised in a single cluster (Pollard and Holland, 2000). In addition, a second *Gsx*-homolog and a second and a third *Cdx* homolog

were found, all in different chromosomal locations in the genomes of both species. This suggests that duplication of the ParaHox cluster has occurred in the lineage leading to the vertebrates, followed by gradual loss of specific members of various clusters, leading to one intact ParaHox cluster in modern day vertebrates (Pollard and Holland, 2000; Holland, 2001). Interestingly, spatial colinearity has also been found for the ParaHox cluster (Brooke *et al.*, 1998), supporting the suggested common origin of the *Hox* and *ParaHox* genes. Therefore, comparing features of *Hox* and *ParaHox* genes could gain more insight into the evolutionary history of their expression regulation and function. A feature shared between the *ParaHox* genes of the *Xlox*/*Pdx1* family and *Hox* genes of paralog groups 1 through 10, is hexapeptide-mediated interaction with PBC family cofactors. Interaction of the ParaHox member Pdx1, orthologous to Amphioxus Xlox, with a PBC class member has been shown to be essential for pancreatic development (Swift *et al.*, 1998; Goudet *et al.*, 1999; Dutta *et al.*, 2001; Liu *et al.*, 2001; Kim *et al.*, 2002). This interaction is mediated via a hexapeptide present in Pdx1 (Peers *et al.*, 1995; Goudet *et al.*, 1999; Liu *et al.*, 2001; Kim *et al.*, 2002). In addition, Hox/PBC-like bipartite sequences have been found in the proximal promoter of the somatostatin gene, a transcriptional target gene of Pdx1 (Goudet *et al.*, 1999). This led us to investigate whether the other *ParaHox* genes might interact with PBC class cofactors. In chapter 6, we report conservation of hexapeptide-flanking sequences of Cdx and Pdx1 proteins, present in a wide range of species, resembling the conservation found in Hox group 1 through 8 proteins (Chapter 4 of this thesis). More generally we searched for the presence of a hexapeptide sequences and conservation of flanking sequences in all of the members of the Antp-class of homeodomain proteins, and found a wide distribution and conservation of flanking sequences. This suggests that interactions between Antp-class homeodomain proteins and TALE-class co-factors occurred early during evolution.

References

Abu-Shaar, M., Ryoo, H. D., and Mann, R. S., (1999), Control of the nuclear localization of Extradenticle by competing nuclear import and export signals. Genes Dev. 13: 935-945.

Acampora, D., D'Esposito, M., Faiella, A., Pannese, M., Migliaccio, E., Morelli, F., Stornaiuolo, A., Nigro, V., Simeone, A., Boncinelli, E., (1989), The human HOX gene family. Nucl. Acids Res. 17: 10385-10402.

Bailey, W. J., Kim, J., Wagner, G. P., Ruddle, F. H., (1997), Phylogenetic reconstruction of vertebrate Hox cluster duplications. Mol. Biol. Evol. 14: 843-853.

Brooke, N. M., Garcia-Fernàndez, J., and Holland, P. W. H., (1998), The ParaHox gene cluster is an evolutionary sister of the Hox gene cluster. Nature 392: 920-922.

Bürglin, T., in A Guidebook for Homeobox Genes (ed Duboule, D.) 25-71, Oxford University Press, Oxford, 1994).

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

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. Proc. Natl. Sci. U S A. 93: 5223-5228.

Chan, S. K., Ryoo, H. D., Gould, A., Krumlauf, R., Mann, R. S., (1997), Switching the *in vivo* specificity of a minimal *Hox*-responsive element. Development 124: 2007-2014.

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. Moll. Cell. Biol. 16: 1734-1745.

Choe, S. K., Vlachakis, N., and Sagerström, C. G., (2001), Meis family proteins are required for hindbrain development in the zebrafish. Development 128: 585-595.

Christian, J. L., Moon, R. T., (1993), Interactions between *Xwnt8* and Spemann organiser signalling pathways generate dorsoventral pattern in the embryonic mesoderm of *Xenopus*. Genes Dev. 7: 13-28.

Cohen, R. N., Brzostek, S., Kim, B., Chorev, M., Wondisford, F. E., and Hollenberg, A. N., (2001), The specificity of interactions between nuclear hormone receptors and corepressors is mediated by distinct amino acid sequences within the interacting domains. Mol. Endocrinol. 15: 1049-1061.

Condie, B. G. and Capecchi, M. R., (1994), Mice with targeted disruptions in the paralogous genes hoxa-3 and hoxd-3 reveal synergistic interactions. Nature 370: 304-307.

Davis, A. P., Witte, D. P., Hsich-Li, H. M., Potter, S. S., and Capecchi, M. R., (1995), Absence of radius and ulna in mice lacking hoxa-11 and hoxd-11. Nature 375: 791-795.

Deschamps, J., van den Akker, E., Forlani, S., de Graaff, W., Oosterveen, T., Roelen, B., and Roelfsema, J., (1999), Initiation, establishment and maintenance of *Hox* gene expression patterns in the mouse. Int. J. Dev. Biol. 43: 635-650.

Di Rocco, G., Mavilio, F., and Zappavigna, V., (1997), Functional dissection of a transcriptional active, target-specific Hox-Pbx complex. EMBO J. 16:3644-3654.

Dibner, C., Elias, S., and Frank, D, (2001), XMeis3 protein is required for proper hindbrain patterning in *Xenopus laevis* embryos. Dev 128: 3415-3426.

van Dijk, M. A. and Murre, C., (1994), extradenticle Raises the DNA binding specificity of Homeotic Selector Gene Products. Cell 78: 617-624.

van Dijk, M. A., Peltenburg, L. T. C., Murre, C., (1995), Hox gene products modulate the DNA binding activity of Pbx1 and Pbx2. Mech. Dev. 52: 99-108.

Duboule, D. and Dollé, P., (1989), The structural and functional organisation of the murine HOX gene family resembles that of Drosophila homeotic genes. EMBO J. 8: 1479-1505.

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

Durston, A. J., van der Wees, J., Pijnappel, W. W., and Godsave, S. F., (1998), Retinoids and related signals in early development of the vertebrate central nervous system. Curr. Top. Dev. Biol. 40: 111-175.

Dutta, S., Gannon, M., Peers, B., Wright, C., Bonner-Weir, S., and Montminy, M., (2001), PDX:PBX complexes are required for normal proliferation of pancreatic cells during development. Proc. Natl. Acad. Sci. U S A. 98: 1065-1070.

Erter, C. E., Wilm, T. P., Basler, N., Wright, C. V. E., Solnica-Krezel, L., (2001), Wnt8 is required in lateral mesodermal precursors for neural posteriorization in vivo. Development 128: 3571-3583.

Ferretti, E., Marshall, H., Pöpperl, H., Maconochie, M., Krumlauf, R., Blasi, F., (2000), Segmental expression of Hoxb2 in r4 requires two separate sites that integrate cooperative interactions between Prep1, Pbx and Hox proteins. Development. 127: 155-66.

Fredieu, J. R., Yanzhen, C., Maier, D., Danilchik, V., Christian, J. L., (1997), Xwnt-8 and Lithium Can Act upon either Dorsal Mesodermal or Neurectodermal Cells to Cause a Loss of Forebrain in *Xenopus* Embryos. Dev. Biol. 186: 100-114.

Gavalas A., Studer, M., Lumsden, A., Rijli, F. M., Krumlauf, R., and Chambon, P., (1998), Hoxa1 and Hoxb1 synergize in patterning the hindbrain, cranial nerves and second pharyngeal arch. Development 125: 1123-1136.

Gavalas, A., and Krumlauf, R., (2000), Retinoid signalling and hindbrain patterning. Curr. Opin. Genet. Dev. 10: 1123-1136.

Garcia-Fernàndez, J., and Holland, P. W. H., (1994), Archetypal organisation of the Amphioxus *Hox* gene cluster. Nature 370: 563-566.

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.

Gaunt, S. J., (1991), Expression patterns of mouse Hox genes: clues to an understanding of developmental and evolutionary strategies. Bioessays 13: 505-513.

Gaunt, S. J., (2000), Evolutionary shifts of vertebrate structures and *Hox* expression up and down the axial series of segments: a consideration of possible mechanisms. Int. J. Dev. Biol. 44: 109-117.

Gaunt, S. J. and Strachan, L. (1996), Temporal colinearity in expression of anterior *Hox* genes in developing chick embryos. Dev. Dyn. 207: 270-280.

Goudet, G., Delhalle, S., Biemar, F., Martial, J. A., and Peers, B., (1999), Functional and Cooperative Interactions between the Homeodomain PDX1, Pbx, and Prep1 Factors on the Somatostatin Promoter. J. Biol. Chem. 274: 4067-4073.

Gould, A., Morrison, A., Sproat, G., White, R. A., and Krumlauf, R., (1997), Positive cross-regulation and enhancer sharing: two mechanisms for specifying overlapping *Hox* expression patterns. Genes Dev. 11: 900-913.

Graham, A., Papalopulu, N., and Krumlauf, R., (1989), The murine and Drosophila homeobox gene complexes have common features of organisation and expression. Cell 57: 367-378.

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

Grieder, N. C., Marty, T., Ryoo, H. D., Mann, R. S., Affolter, M., (1997), Synergistic activation of a Drosophila enhancer by HOM/EXD and DPP signaling. EMBO J. 16: 7402-7410.

Hofmann, C., and Eichele, G., (1994), in The Retinoids: Biology, Chemistry and Medicine (eds. M. B. Sporn, A. B. Roberts and D. S. Goodman) 387-441 (Raven Press, New York).

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

Hayashi, S., and Scott, M. P., (1990), What determines the specificity of action of Drosophila homeodomain proteins? Cell 63: 883-894.

Holland, P. W. H., (2001), Beyond the Hox: how widespread is homeobox gene clustering? J. Anat. 199: 13-23.

Hughes, C. L., and Kaufman T. C., (2002), Exploring the myriapod body plan: expression patterns of the ten Hox genes in a centipede. Development 129: 1225-1238.

Izpisua-Belmonte, J. C., Falkenstein, H., Dollé, P., Renucci, A., and Duboule, D., (1991), Murine genes related to the Drosophila AbdB homeotic genes are sequentially expressed during development of the posterior part of the body. EMBO J. 350: 585-589.

Jabet, C., Gitti, R., Summers, M. F., and Wolberger, C., (1999), NMR Studies of the Pbx1 TALE Homeodomain Protein Free in Solution and Bound to DNA: Proposal for a Mechanism of HoxB1-Pbx1-DNA Complex Assembly. J. Mol. Biol. 291: 521-530.

Jaw, T. J., You, L. R., Knoepfler, P. S., Yao, L. C., Pai, C. Y., Tang, C. Y., Chang, L. P., Berthelsen, J., Blasi, F., Kamps, M. P., Sun, Y. H., (2000), Direct interaction of two homeoproteins, Homothorax and Extradenticle, is essential for EXD nuclear localization and function. Mech. Dev. 91: 279-291.

Jorgensen, E. M., and Garber, R. L., (1987), Function and misfunction of the two promoters of the *Drosophila Antennapedia* gene. Genes and Dev. 1: 544-555.

Kessel, M., and Gruss, P., (1990), Murine developmental control genes. Science 249: 347-379.

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.

Kiecker, C. and Niehrs, C., (2001), A morphogen gradient of Wnt/β-catenin signalling regulates anteroposterior neural patterning in *Xenopus*. Development 128: 4189-4201.

Kim, C. B., Amemiya, A., Bailey, K., Kawasaki, J., Mezey, J., Miller, W., Minoshima, S., Shimizu, N., Wagner, G., Ruddle, F., (2000), Hox cluster

genomics in the horn shark, Heterodontus francisci. Proc. Natl. Acad. Sci. U S A. 97: 1655-1660.

Kim, S. K., Selleri, L., Lee, J. S., Zhang, A. Y., Gu, X., Jacobs, Y., Cleary, M. L., 2002, Pbx1 inactivation disrupts pancreas development and in Ipf1-deficient mice promotes diabetes mellitus. Nat. Genet. 30(4): 430-435.

Knoepfler, P. S., and Kamps, M. P., (1995), The Pentapeptide Motif of Hox Proteins Is Required for Cooperative DNA Binding with Pbx1, Physically Contacts Pbx1, and Enhances DNA binding by Pbx1. Moll. Cell. Biol. 15: 5811-5819.

Knoepfler, P. S., Bergstrom, D. A., Uetsuki, T., Korytko, D., Sun, Y. H., Wright, W. E., Tapscott, S. J., and Kamps, M. P., (1999), A conserved motif N-terminal to the DNA-binding domains of myogenic bHLH transcription factors mediates cooperative DNA binding with Pbx-Meis1/Prep1. Nucl. Acid Res. 27: 3752-3761.

Knoepfler, P. S., Lu, Q., and Kamps, M. P., (1996), Pbx1-Hox heterodimers bind DNA on inseparable half-sites that permit intrinsic DNA binding specificity of the Hox partner at nucleotides 3' to a TAAT motif. Nucl. Acid Res. 24: 2288-2294.

Laughon, A., (1991), DNA binding specificity of homeodomains. Biochemistry 30: 11357-11367.

Liu, Y., MacDonald, R. J., and Swift, G. H., (2001), DNA Binding and Transcriptional Activation by a PDX1-PBX1b-MEIS2b Trimer and Cooperation with a Pancreas-specific Basic Helix-Loop-Helix Complex. J. Biol. Chem. 276: 17985-17993.

Lufkin, T., Mark, M., Hart, C. P., Dollé, P., LeMeur, M., Chambon, P., (1992), Homeotic transformation of the occipital bones of the skull by ectopic expression of a homeobox gene. Nature 359: 835- 841.

Maconochie, M. K., Nonchev, S., Studer, M., Chan, S. K., Pöpperl, H., Sham, M. H., Mann, R. S., and Krumlauf, R., (1997), Cross-regulation in the mouse *HoxB* complex: the expression of *Hoxb2* in rhombomere 4 is regulated by *Hoxb1*. Genes Dev. 11: 1885-1895.

Mann, R. S. and Affolter, M., (1998), Hox proteins meet more partners. Curr. Opin. Genet. Dev. 8: 7402-7410.

Mann, R. S. and Chan, S. K., (1996), Extra specificity from extradenticle: the partnership between HOX and PBX/EXD homeodomain proteins. Trends Genet. 12: 258-262.

McGinnis, W. and Krumlauf, R., (1992), Homeobox Genes and Axial Patterning. Cell 68: 283-302.

Passner, J. M., Ryoo, H. D., Shen, L., Mann, R. S., and Aggarwal, A. K., (1999), Structure of a DNA-bound Ultrabithorax-Extradenticle homeodomain complex. Nature 397: 714-718.

Peers, B., Sharma, S., Johnson, T., Kamps, M., Montminy, M., (1995), The pancreatic islet factor STF-1 binds cooperatively with Pbx to a regulatory element in the somatostatin promoter: importance of the FPWMK motif and of the homeodomain. Mol. Cell Biol. 15: 7091-7097.

Peltenburg, L. T. C. and Murre, C., (1996), Engrailed and Hox homeodomain proteins contain a related Pbx interaction motif that recognizes a common structure present in Pbx. EMBO J. 15: 3385-3393.

Phelan, M. L. and Featherstone, M. S., (1997), Distinct HOX N-terminal Arm Residues Are Responsible for Specificity of DNA Recognition by HOX Monomers and HOX-PBX Heterodimers. J. Biol. Chem. 272: 8635-8643.

Phelan, M. L., Rambaldi, I., and Featherstone, M. S., (1995), Cooperative interactions between Hox and PBX proteins mediated by a conserved peptide motif. Mol. Cell. Biol. 15: 3989-3997.

Piper, D. E., Batchelor, A. H., Chang, C. P., Cleary, M. L., Wolberger, C., (1999), Structure of a HoxB1-Pbx1 Heterodimer Bound to DNA: Role of the Hexapeptide and a Fourth Homeodomain Helix in Complex Formation. Cell 96: 587-597.

Pollard, S. L. and Holland, P. W. H., (2000), Evidence for 14 homeobox gene clusters in human genome ancestry. Curr. Biol. 10: 1059-1062.

Pöpperl, H., Bienz, M., Struder, M., Chan, S. K., Aparicio, S., Brenner, S., Mann, R. S., and Krumlauf, R., (1995), Segmental expression of *Hoxb1* is Controlled by a Highly conserved Autoregulatory Loop Dependent upon *exd*/*pbx*. Cell 81: 1031-1042.

Rieckhof, G. E., Casares, F., Ryoo, H. D., Abu-Shaar, M., Mann, R. S., (1997). Nuclear Translocation of Extradenticle Requires *homothora*x, which Encodes an Extradenticle-Related Homeodomain Protein. Cell 91: 171-183.

Rossel, M. and Capecchi, M. R., (1999), Mice mutant for both Hoxa1 and Hoxb1 show extensive remodelling of the hindbrain and defects in craniofacial development. Development 126: 5027-5040.

Ruddle F. H., Bentley, M., Murtha, M. T., Risch, N., (1994) Gene loss and gain in the evolution of the vertebrates. Development (Supplement): 155-161.

Ruvkun, G., and Hobert, O., (1998), The Taxonomy of Developmental Control in *Caenorhabditis elegans*. Science 282: 2033-2041.

Ryoo, H. D., and Mann, R. S., (1999), The control of trunk Hox specificity and activity by Extradenticle. Genes & Dev. 13: 1704-1716.

Ryoo, H. D., Marty, T., Casares, F., Affolter, M., Mann, R. S., (1999) Regulation of Hox target genes by a DNA bound Homothorax/Hox/Extradenticle complex. Development 126: 5137-5148.

Salzberg, A., Elias, S., Nachaliel, N., Bonstein, L., Henig, C., Frank, D., (1999), A Meis family member protein caudalizes neural cell fates in *Xenopus*. Mech. Dev. 80: 3-13.

Schneuwly, S., Kuroiwa, A., Baumgartner, P., and Gehring, W. J., (1986), Structural organization and sequence of the homeotic gene *Antennapedia* of *Drosophila melanogaster*. EMBO J. 5: 733-739.

Schughart, K., Kappen, C., Ruddle F. H., (1989), Duplication of large genomic regions during the evolution of vertebrate homeobox genes. Proc. Natl. Acad. Sci. U S A 86: 7067-7071.

Shen, W. F., Chang, C. P., Rozenfeld, S., Sauvageau, G., Humphries, R. K., Lu, M., Lawrence, H. J., Cleary, M. L. and Largman, C., (1996), Hox

homeodomain proteins exhibit selective complex stabilities with Pbx and DNA. Nucl. Acid Res. 24: 898-906.

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.

Studer, M., Gavalas, A., Marshall, H., Ariza-McNaughton, L., Rijli, F. M., Chambon, P., Krumlauf, R., (1998), Genetic interactions between Hoxa1 and Hoxb1 reveal new roles in regulation of early hindbrain patterning. Development 125: 1025-1036.

Swift, G. H., Liu, Y., Rose, S. D., Bischof, L. J., Steelman, S., Buchberg, A. M., Wright, V. E., and MacDonald, R. J., (1998), An endocrine-Exocrine Switch in the activity of the pancreatic Homeodomain Protein PDX1 through Formation of a Trimeric complex with PBX1b and MRG1 (MEIS). Moll. Cell. Biol. 18: 5109-5120.

Vlachakis, N., Choe, S. K., and Sagerström, C. G., (2001), Meis3 synergizes with Pbx4 and Hoxb1b in promoting hindbrain fates in the zebrafish. Development 128: 1299-1312.

Wacker, S. A., Jansen, H. J., McNulty, C. L., Houtzager, E., and Durston, A. J., Timed interactions between a Hox cascade and the Spemann organiser generate positional information during vertebrate gastrulation. Submitted.

Waskiewicz, A. J., Rikhof, H. A., Hernandez, R. E., and Moens, C. B., (2001), Zebrafish Meis functions to stabilize Pbx proteins and regulate hindbrain patterning. Development 128: 4139-4151.

Zákány, J., Gérard, M., Favier, B., Potter, S. S., and Duboule, D., (1996), Functional equivalence and rescue among group 11 *Hox* gene products in vertebral patterning. Dev. Biol. 176: 325-328.

Chapter 2

Xwnt8 directly initiates expression of the most anterior *Hox* gene

Paul M. J. In der Rieden and Antony J. Durston

Hubrecht Laboratorium,
Nederlands Instituut voor Ontwikkelingsbiologie,
Uppsalalaan 8, 3584 CT, Utrecht, The Netherlands.

Abstract

Hox transcription factors serve an essential role in patterning the anteroposterior axis, during embryogenesis. Recently, an early expression sequence of Hox genes was found in mesoderm of *Xenopus* gastrulae, this led us to investigate whether factors known to pattern the mesoderm are involved in initiation of *Hox* gene expression. Here we present evidence that Xwnt8 is necessary for initiation of *Hoxd1* expression, and that *Hoxd1* is a direct target of Tcf/Lef signalling, during gastrulation. In addition, Xwnt8 loss- and gain-of-function leads to an induction of *Hoxc6* expression; this sheds new light on the proposed wnt gradient in patterning the *Xenopus* central nervous system (Kiecker and Niehrs, 2001).

Introduction

Hox proteins are involved in the specification of positional identities along the anteroposterior (AP) and other embryonic axis in a wide range of animal species, including vertebrates (Bürglin *et al.*, 1991; McGinnis and Krumlauf, 1992; Bürglin and Ruvkun, 1993; Lawrence and Morata, 1994; Manak and Scott, 1994). *Hox* genes are organised in four clusters located on different chromosomes. The clusters are thought to have arisen by tandem duplication of a single gene, followed, in vertebrates, by duplication of the cluster itself (Schughart *et al.*, 1989; Ruddle *et al.*, 1994; Bailey *et al.*, 1997; Greer *et al.*, 2000). As a consequence, *Hox* genes occupying the same relative position along the 5' to 3' chromosomal coordinate, named paralogous genes, share more similarity in sequence and expression pattern than do adjacent *Hox* genes on the same chromosome (Greer *et al.*, 2000). A phenomenon of particular interest is that *Hox* genes located at the 3’ end of a cluster are expressed earlier and more anteriorly than the subsequently more 5’ located genes (Gaunt *et al.*, 1988; Duboule, 1994; Gaunt and Strachan, 1996). How spatiotemporal colinearity in the expression of *Hox* genes is regulated is intriguing but, to date, not well understood.
Recently, the early *Hox* expression patterns have been analysed in *Xenopus* gastrula embryos (Wacker *et al*., submitted). This revealed spatiotemporally colinear initiation of expression of a sequence of *Hox* genes within a horseshoe-shaped domain in ventrolateral marginal zone mesoderm at different stages during gastrulation, followed by sequential dorsalisation of each *Hox* expression zone into a stable AP zone in axial mesoderm and the neural plate. To gain further insight into the regulation of spatiotemporal *Hox*

expression in marginal zone mesoderm, we wished to identify factors involved the initiation of mesodermal *Hox* expression.

In *Xenopus*, *Xwnt8* expression is first detected in late blastula stage embryos. Expression is found in all cells of the marginal zone with the exception of the cells centred on the dorsal midline. This pattern of expression in the ventrolateral marginal zone persists during gastrulation (Christian and Moon, 1993). Ectopic expression of *Xwnt8* posteriorises neurectoderm (Fredieu *et al.*, 1997; Erter *et al.*, 2001; Kiecker and Niehrs, 2001), a feature also known for *Hox* genes. Conversely, gain-of-function for *Xdkk1*, a secreted Wnt antagonist (Glinka *et al.*, 1998), downregulates the expression of *Hoxd1* in neurectoderm of *Xenopus* embryos (Kiecker and Niehrs, 2001). In mouse and chick embryos, the expression patterns of the *Xwnt8* orthologs are indicative for a possible function in the regulation of expression of labial-type *Hox* genes. In chick embryos, the expression of *Cwnt8C* immediately precedes the localisation of *Hoxb1* expression to rhombomere 4 (Hume and Dodd, 1993). In mouse embryos, expression of *Mwnt8* is found in the presumptive rhombomere 4 region (Bouillet *et al.*, 1996). In *Caenorhabditis elegans*, Wnt/WG signalling elements are involved in the regulation of *ceh-13*, the labial ortholog of the worm (Streit *et al.*, 2002). These properties makes *Xwnt8* is a good candidate to fulfil the role of initiator of *Hox* expression in marginal zone mesoderm of *Xenopus* embryos.

Xwnt8 is a member of the *Wnt* family of secreted glycoproteins, which act as ligands, activating receptor-mediated signal transduction pathways (reviewed in Moon *et al.*, 2002 and references therein). After binding of *Xwnt8* to suitable receptors, intracellular signals are transduced by the canonical Wnt pathway (Darken and Wilson, 2001), which acts through a rise in cytosolic and subsequent nuclear levels of β-catenin, influencing the function of Tcf/Lef transcription factors. Misexpression of synthetic *Xwnt8* mRNA on the presumptive ventral side, before the activation of the zygotic genome, leads to formation of a secondary axis (Sokol *et al.*, 1991), while later activation of *Xwnt8* expression leads to posteriorisation of the primary axis (Christian and Moon, 1993). In *Xenopus* embryos, it has been shown that β-catenin induced axis formation is mediated via the transcription factor *XTcf3* (Molenaar *et al.*, 1996). The early and late effects of ectopic *Xwnt8* on axis formation can be mimicked by timed activation of an activated form of XTcf3 (Darken and Wilson, 2001).

So, is *Xwnt8* signalling involved in the initiation of mesodermal *Hox* expression? To answer this question we employed the following strategies. First, we made a detailed description of the early expression patterns of *Xwnt8* and examined whether the expression of *Xwnt8* coincides with the

expression of *Hoxd1*, *Hoxb4*, and *Hoxc6* during gastrulation. These three genes were chosen because they are expressed in well-defined spatial domains in the neurectoderm, corresponding to the identities of rhombomeres 4 and 5 (*Hoxd1*, Kolm and Sive, 1995b), posterior hindbrain (*Hoxb4*, Harvey and Melton, 1988), and anterior spinal cord (*Hoxc6*, Oliver *et al.*, 1988; De Robertis *et al.*, 1989), in addition, the spatiotemporal colinear expression of these genes in ventrolateral mesoderm has been described (Wacker *et al.*, submitted). We report a significant overlap in expression of *Xwnt8* and the assayed *Hox* genes in ventrolateral mesoderm during gastrula stages.
Next, we analysed the effects of *Xwnt8* loss-of-function, using a morpholino-based strategy, on development and the expression of early *Hox* genes during gastrulation. Xwnt8 loss-of-function leads to an anteriorisation of embryos, accompanied by a reduction in expression of *Hoxd1*, but not of *Hoxb4*.
Next, we performed *Xwnt8* gain-of-function experiments. This results in posteriorised embryos and a strong upregulation of *Hoxd1* expression. To investigate whether the observed effects on *Hox* expression by Wnt8 gain-of-function are direct, we undertook an approach using a fusion of an activated form of XTcf3 to the ligand-binding domain of the glucocorticoid receptor, which allows hormonal regulation of nuclear translocation. Activation of a dominant positive form of XTcf3, shortly before gastrulation, leads in a direct fashion to upregulation of *Hoxd1* expression. In this report, we present evidence that *Xwnt8* function is necessary and sufficient to induce the expression of *Hoxd1* in mesoderm and ectoderm of gastrula embryos. Furthermore we show that initiation of *Hoxd1* expression can be performed, in a direct fashion, by Tcf/Lef signalling.

Results

Xwnt8 and anterior *Hox* genes have partially overlapping expression domains during gastrulation

If *Xwnt8* is involved in the initiation of *Hox* gene expression, it needs to be co-expressed with *Hox* genes. Because Wnt family members are secreted factors their functional domains could extend beyond the borders of their expression domains, but nonetheless overlapping expression of *Xwnt8* and *Hox* genes could reveal functional relations. We compared the detailed expression patterns of *Xwnt8* and three early *Hox* genes in gastrula and early neurula stage embryos. Early during gastrulation *Xwnt8* is expressed in a horseshoe-like pattern in the mesoderm, showing a gap of expression in the organiser mesoderm (Fig. 1A). During progression of gastrulation *Xwnt8*

expression expands in animal direction (Fig. 1B and 1C). Expression of *Xwnt8* is lost at the ventralmost side of the embryo around stage 12 (Fig. 1C). Expression is maintained in dorsolateral mesodermal domains close to the blastopore, and in involuted mesoderm (Fig. 1C). During early neurulation, three domains of *Xwnt8* expression can be observed on either side of the midline: a domain in the paraxial mesoderm, a domain in presumptive hindbrain neurectoderm with anteriormost expression overlying the anterior expression domain in paraxial mesoderm, and a posterior domain in dorsolateral mesoderm (Fig. 1E). Expression of *Hoxd1* starts in a horseshoe-like pattern in marginal zone mesoderm at stage 10.25 (Fig. 2B) and as gastrulation progresses two dorsolateral located domains become prominent (Fig. 2B). At stage 11.5 the ectoderm overlying the dorsolateral mesodermal expression domains starts to express *Hoxd1* (Fig. 2B). Early during neurulation expression of *Hoxd1* can be found anteriorly in ectoderm, in lateral mesoderm extending backwards to the almost closed blastopore (Fig. 2B). The expression patterns of *Hoxd1* and *Xwnt8* in gastrula stages show a clear overlap (compare Fig. 2A to 2B). During early gastrulation the overlap can be found in marginal zone mesoderm. At stage 13, expression of both genes is found in neurectoderm, in paraxial and posterior mesoderm expression of *Xwnt*8 confides within the domain of *Hoxd1* expression (compare Fig. 2A to 2B). Initiation of *Hoxb4* expression takes place later during gastrulation than the initiation of *Hoxd1* expression (stage 10.5) but in a similar, nested domain in marginal zone mesoderm (Fig. 2C). At stage 12, ectoderm, overlying the dorsolateral mesodermal expression domains, starts to express *Hoxb4* (Fig. 2C). At stage, 13 this ectodermal expression of *Hoxb4* is located posteriorly to the ectodermal expression of *Hoxd1* (compare Fig. 2B to 2C). Expression of *Hoxb4* overlaps with that of *Xwnt8* in marginal zone mesoderm (compare Fig. 2A and 2C). At stage 12 this overlap is restricted to the dorsolateral domain of *Hoxb4* expression, at stage 13 both *Xwnt8* and *Hoxb4* are co-expressed in paraxial mesoderm, while no overlap in neurectoderm can be observed. Expression of *Hoxc6* is initiated in a similar pattern to that of *Hoxd1* and *Hoxb4*, starting at stage 11.5 in marginal zone mesoderm (Fig. 2D). At stage 13 expression of *Hoxc6* in ectoderm overlying the dorsolateral mesodermal expression domain can be observed (Fig. 2D) with an anterior border of expression in a more posterior position than the anterior border of *Hoxb4* expression (compare Fig. 2C to 2D). The expression patterns of *Hoxc6* and *Xwnt8* overlap in marginal zone mesoderm but not in neurectoderm. Later during gastrulation, the expression overlap in posterior dorsolateral mesoderm remains during these late gastrula stages

(compare Fig. 2D to 2A). These results show that *Xwnt8* could serve a role as an initiator of *Hox* gene expression during gastrulation.

Xwnt8 loss-of-function leads to anteriorisation of embryos and loss of *Hoxd1* expression

To investigate whether *Xwnt8* is of importance for the expression initiation of *Hoxd-1*, *Hoxb-4*, and *Hoxc-6* and we employed a loss-of-function method using a *Xwnt8* morpholino antisense oligonucleotide (MO^{Xwnt8}). A number of loss-of-function strategies have been used to study the function of Xwnt8: *dnWnt8* (Hoppler *et al.*, 1996), *Xdkk-1* (Glinka *et al.*, 1998), and *Sizzled* (Salic *et al.*, 1997). The advantage of a morpholino-based approach is the reported high specificity (reviewed in Heasman, 2002 and references therein). By binding of the morpholino, to sequences overlapping, or lying adjacent to, the start site of translation, the targeted mRNA is not translated (reviewed in Heasman, 2002). This results in a potentially more specific Xwnt8 loss-of-function method compared to overexpressing antimorphic forms of *Xwnt8* or Wnt antagonists.
MO^{Xwnt8} was injected into the animal hemisphere of embryos at the one-cell stage, resulting in spreading of the MO^{Xwnt8} all over the embryo, subsequently, the embryos were allowed to develop until control embryos reached stage 24 (Fig. 3A) or stage 35 (Fig. 3B). Knocking down Xwnt8 function by injection of MO^{Xwnt8} leads to anteriorisation of the embryo in a concentration dependent manner (Fig. 3). In MO^{Xwnt8} injected embryos the axis was reduced, and an enlargement of the cement gland was observed (compare Fig. 3A to 3C and 3D, and Fig. 3B to 3E). This phenotype has also been reported for the other mentioned Xwnt8 (or Wnt), loss-of-function methods: *dnWnt8* (Hoppler *et al.*, 1996), *Xdkk-1* (Glinka *et al.*, 1998), *Sizzled* (Salic *et al.*, 1997). In zebrafish embryos, injection of morpholinos directed against both the *Zwnt8* ORFs found (Erter *et al.*, 2001; Lekven *et al.*, 2001) leads to comparable effects on development of the embryos as observed for *Xenopus* using the MO^{Xwnt8}. A control morpholino (MO^{contr}), in sequence unrelated to MO^{Xwnt8}, was injected in the same amounts as the MO^{Xwnt8}. Abnormalities in the development of embryos injected with MO^{contr} were not observed (data not shown). The specificity of the MO^{Xwnt8} was further shown by rescue of the Xwnt8 loss-of-function phenotype with CS2-*Xwnt8* plasmid DNA (see materials and methods for details). 64 ng of MO^{Xwnt8} and 20 pg CS2-*Xwnt8* were injected either singly or in combination into the animal hemisphere of embryos at the one-cell stage. The embryos receiving both the

MO^{Xwnt8} and the CS2-*Xwnt8* show a clear reduction in size of the cement gland as compared to the single injection of the MO^{Xwnt8} (Fig. 3F).
After confirming that the MO^{Xwnt8} is a valid Xwnt8 loss-of-function strategy we investigated its effects on the expression patterns of *Hoxd1*, *Hoxb4*, and *Hoxc6*. Mesodermal expression of *Hoxd1* was strongly downregulated, and the distance between the two dorsolateral domains of expression in marginal zone mesoderm was increased by Xwnt8 loss-of-function (Fig. 4A). Ectodermal expression of *Hoxd1* is also downregulated in injected embryos (Fig. 4A). Expression of *Hoxb4* in mesoderm and ectoderm appeared unaltered by Xwnt8 loss-of-function (Fig. 4B). Expression of *Hoxc6* was ectopically upregulated in dorsal mesoderm of stage 10.5 embryos and in mesoderm and dorsal ectoderm of embryos at stage 11.5 (Fig. 4C). *In situ* hybridisations were performed on embryos injected with 64 ng of MO^{contr}, for all markers studied this results in unaltered expression (data not shown).

Misexpression of *Xwnt8* after mid-blastula transition leads to an upregulation of expression of *Hoxd1* and *Hoxc6*, but not of *Hoxb4*

To study the effects of *Xwnt8* gain-of-function on gastrulation and neurulation we generated a construct driving expression of *Xwnt8* after the mid blastula transition (MBT). This avoids the early, dorsalising, activity found following *Xwnt8* synthetic mRNA injections (Smith and Harland, 1991; Sokol *et al.*, 1991). To this end, we generated a plasmid containing the full-length coding region of *Xwnt8* in the CS2+ vector (Rupp *et al.*, 1994) and named the construct CS2-*Xwnt8*. The CS2+ vector harbours a sCMV promoter leading to efficient expression and subsequent translation of the derived mRNA in *Xenopus* embryos after MBT (Turner and Weintraub, 1994; Kühl *et al.*, 1996). Embryos at the one-cell stage were injected into the animal hemisphere with 100 pg of CS2-*Xwnt8* plasmid, this results in a clear posteriorisation (data not shown). Next, we assayed for the expression of early *Hox* genes, the mesodermal marker *Xbra*, and the posterior marker *Xcad3* in CS2-*Xwnt8* injected embryos. Strong upregulation of the expression of *Hoxd1* could be observed (Fig. 5A). Not only is the expression domain larger as compared to control embryos, but expression also appears earlier and expression can be observed in the organiser field (Fig. 5A). Later during gastrulation ectopic *Hoxd1* expression continues to be present in ectoderm and mesoderm on the midline of the embryo, and is expanded in anterior direction (Fig. 5A). Expression of *Hoxb4* appeared unaltered by Xwnt8 gain-of-function (Fig. 5B). Expression of *Hoxc6* was upregulated, in mesoderm and ectoderm (Fig 5C). This ectopic expression is much earlier than

endogenous *Hoxc6* expression, and is found in dorsal mesoderm, tissue that normally does not express *Hox* genes. Later in gastrulation an expansion of the endogenous horseshoe-shaped domain is observed (data not shown). In early neurula stages, an anterior expansion of the expression of *Hoxc6* in neurectoderm and axial mesoderm is observed (Fig. 5C). Expression of the mesodermal marker *Xbra* was unaltered by the Xwnt8 gain-of-function (Fig. 5D), suggesting that formation of mesoderm was not affected by exogenous *Xwnt8*. Expression of the posterior marker *Xcad3* was upregulated in mesoderm (data not shown) and ectoderm of injected embryos (Fig. 5E), confirming the posteriorising nature on neurectoderm of CS2-*Xwnt8* injection. The different effects that the misexpression of *Xwnt8* has on the expression of *Hoxd1*, *Hoxb4*, and *Hoxc6* demonstrates the complex and dynamic nature of expression regulation of *Hox* genes in marginal zone mesoderm, and suggests a role for *Xwnt8* in this regulation.

Labial type *Hox* genes are direct targets of canonical Wnt signalling

It is has been shown that *Xwnt8* employs the canonical Wnt pathway before and after the onset of gastrulation, stabilizing cytosolic β-catenin and activating gene expression through Tcf/Lef transcription factors (Darken and Wilson, 2001). To investigate whether the induction of anterior *Hox* genes by Xwnt8 signalling is direct we made use of an activated, hormone inducible form of XTcf3, TVGR (Darken and Wilson, 2001). Embryos were injected into the animal hemisphere at the one-cell stage with 100 pg of TVGR. CHX was added at stage 9.5, followed half an hour later by addition of DEX to the appropriate samples, well before the onset of gastrulation and the initiation of *Hox* gene expression. At stage 11 the embryos were harvested for RT-PCR. The results are shown in figure 6. Because CHX was added before the onset of gastrulation, induction of *Hoxd1*, *Hoxb4* and *Hoxc6* expression in control embryos is reduced or absent (Fig. 6). Expression of *Hoxd1* is directly activated by the TVGR, while the expression of *Hoxb4* and *Hoxc6* is slightly upregulated when DEX was added in absence of CHX, demonstrating that induction of expression of *Hoxb4* and *Hoxc6* is indirect.

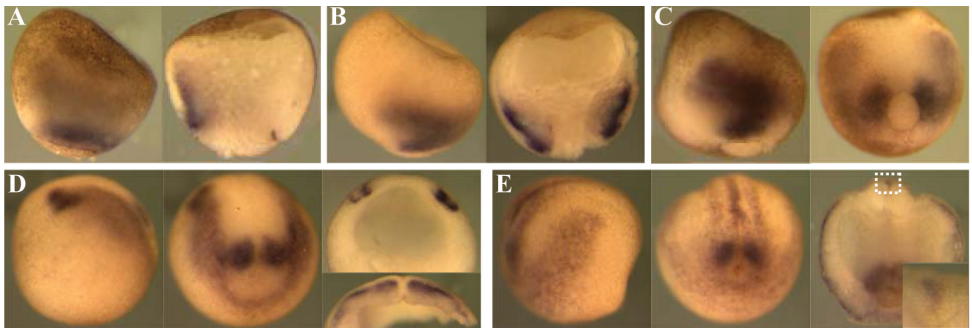

Figure 1. Expression of *Xwnt8* during gastrula and early neurula stages. Embryos were assayed for expression of *Xwnt8* by whole-mount *in situ* hybridisation. In each panel a single embryo is shown. **(A)** Stage 11 embryo, lateral view with dorsal to the right, and a dorsal-to-ventral section of the embryo. *Xwnt8* expression is detected in the ventral and lateral marginal zone mesoderm. **(B)** Stage 11.5 embryo, lateral view with dorsal to the right, and a lateral-to-lateral section. Expression can be found close to the blastopore and in involuted mesoderm. **(C)** Stage 12 embryo, lateral view with dorsal to the right, and a posterior view. Expression of *Xwnt8* can be found in presumptive paraxial mesoderm and expression close to the blastopore is further restricted to dorsolateral positions. **(D)** Stage 13 embryo, lateral view with anterior to the left, a posterior view of the embryo, and two dorsal-to-ventral sections. In the section on the right top of the panel *Xwnt8* expression in presumptive hindbrain is shown, this corresponds to the anteriormost expression in the lateral view. Expression in mesoderm close to the closing blastopore is shown in the bottom right section of the panel and corresponds to expression shown in posterior view. **(E)** Stage 17 embryo, lateral view with anterior to the left, posterior view, and a dorsal-to-ventral section. The anterior ectodermal expression domain, the paraxial expression, and the dorsolateral expression in the mesoderm remain, while a lateral expression domain appears in the ectoderm. In the dorsal-to-ventral section, and in an enlargement on the bottom right of the panel, initiation of *Xwnt8* expression in the neural tube can be found.

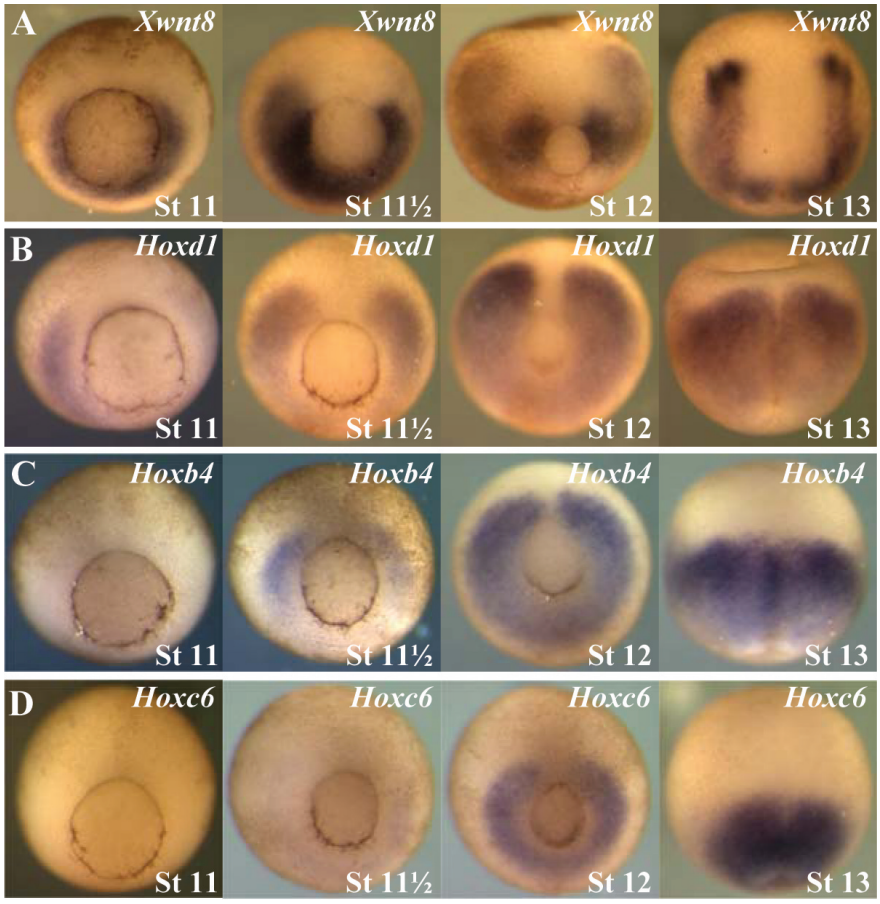

Figure 2. Expression of *Xwnt8*, *Hoxd1*, *Hoxb4*, and *Hoxc6* during gastrulation. Embryos were analysed by whole-mount *in situ* hybridisation for expression of *Xwnt8* **(A)**, *Hoxd1* **(B)**, *Hoxb4* **(C)**, and *Hoxc6* **(D)**. Embryos are shown, going from left to right through the panels, at stage 11, stage 11.5, stage 12 (vegetal views with dorsal up), and at stage 13 (dorsal views with anterior up). *Xwnt8* expression overlaps with the expression of *Hoxd1* in the ventrolateral mesoderm during early gastrulation. At stage 12 the posterior most expression of *Xwnt8* becomes restricted to dorsolateral marginal zone, overlapping with the expression domain of *Hoxd1*. When gastrulation is nearly completed an overlap in expression of *Xwnt8* and *Hoxd1* can be observed in presumptive hindbrain, and paraxial mesoderm. *Hoxb4* and *Xwnt8* show an overlap in their expression patterns during stage 11.5, at stage 12 ectodermal expression of *Hoxb4* is initiated in overlapping the dorsolateral *Xwnt8* expression domain. During late gastrulation an overlap in expression of *Hoxb4* and *Xwnt8* is observed in paraxial mesoderm. Expression of *Hoxc6*, on the other hand, is initiated after the retraction of the *Xwnt8* expression to the dorsolateral domains, therefore an overlap in expression is only observed there. This overlap is still visible at the end of gastrulation. Likewise for *Hoxd1* and *Hoxb4* the first ectodermal expression is initiated of *Hoxc6* found in the ectoderm overlying the posterior dorsolateral domains of *Xwnt8* expression.

Discussion

Ectopic *Xwnt8* directly initiates expression of *Hoxd1*

We report that ectopic *Xwnt8* is sufficient to initiate *Hoxd1* expression in mesoderm and ectoderm of gastrula stage embryos. *Xwnt8* is able to ectopically induce expression of the earliest expressed *Hox* gene in *Xenopus*, *Hoxd1*, and can do so earlier than initiation of endogenous expression and in tissues normally not expressing *Hox* genes as well as in endogenously expressing tissues. Kiecker and Niehrs (2001) have reported that the injection of pCSKA-*Xwnt8* (CSKA-X8, Christian and Moon, 1993) into *Xenopus* embryos does not alter the expression of *Hoxd1*, an apparent contradiction to the results shown in this report, because we do observe an increased expression of *Hoxd1* by ectopic *Xwnt8*. In our hands the pCSKA-*Xwnt8* construct was also not able to initiate the expression of *Hoxd1* in mesoderm or ectoderm. This could be due to the UTR sequences contained in the pCSKA-*Xwnt8* plasmid, as UTR sequences are know to affect the stability of mRNA and to regulate the translation of the messenger (reviewed in Derrigo *et al.*, 2000, and references therein). Our results show that *Xwnt8* is capable of initiating the expression of *Hoxd1*.

Endogenous *Xwnt8* signalling via *Tcf*/*Lef* is necessary for induction of *Hoxd1* in dorsolateral mesoderm and neurectoderm

The necessity of *Xwnt8* function for *Hoxd1* expression in marginal zone mesoderm is shown by *Xwnt8* loss-of-function experiments, where a strong reduction in *Hoxd1* expression can be observed. The expression of *Hoxb4* appears largely unaffected by *Xwnt8* loss-of-function. We conclude that *Hoxb4* is downstream of a different parallel input. A mechanism whereby different inputs are capable of starting *Hox* expression from different *Hox* paralog groups in the cluster could be of importance in the regulation of the *Hox* gene expression and thus for patterning the anteroposterior axis. Striking is the effect on expression of *Hoxc6* by Xwnt8 loss-of-function. *Hoxc6* is up regulated in dorsal mesoderm and ectoderm, tissues normally not expressing *Hoxc6*, significantly earlier than endogenous expression. An explanation could be that *Xwnt8* loss-of-function leads to a reduction of hindbrain identity in the dorsal ectoderm, by reduction of *Hoxd1* expression, resulting in posteriorisation to spinal cord identity, as shown by expression of *Hoxc6*. The necessity of Wnt signalling for the expression of labial-type *Hox* genes is supported by recent findings in *C. elegans* (Streit *et al.*, 2002). It was found

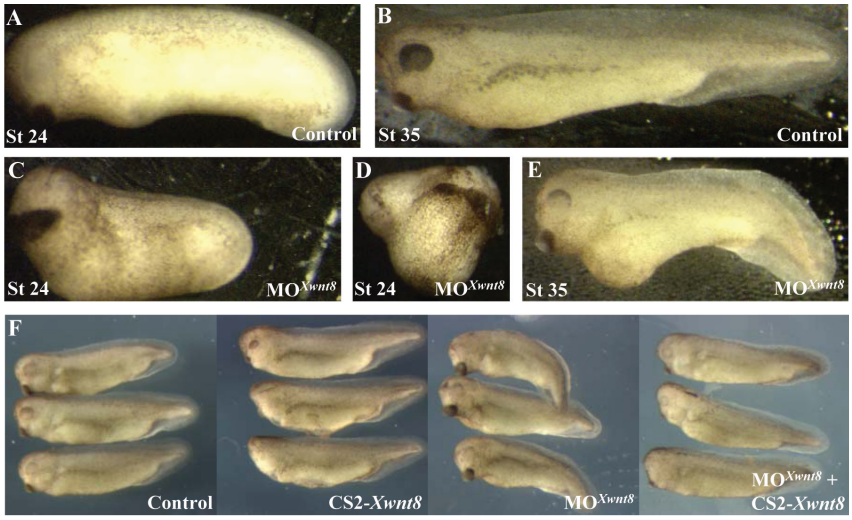

Figure 3. Effects of Xwnt8 loss-of-function on phenotype and rescue of MO^{Xwnt8}. Embryos at the one-cell stage were injected into the animal hemisphere with 64 ng of MO^{Xwnt8}, and allowed to develop until the control embryos reached stage 24 **(A)** or stage 35 **(B)**. In the majority of the embryos the axis is reduced and the head is enlarged, as well as the anterior most structure, the cement gland **(C), (D),**.and **(E)**. **(F)** The specificity of the MO^{Xwnt8} is shown by rescue with CS2-*Xwnt8* plasmid. Embryos were injected with 20 pg CS2-*Xwnt8,* 64 ng MO^{Xwnt8} or with both and compared to control embryos.

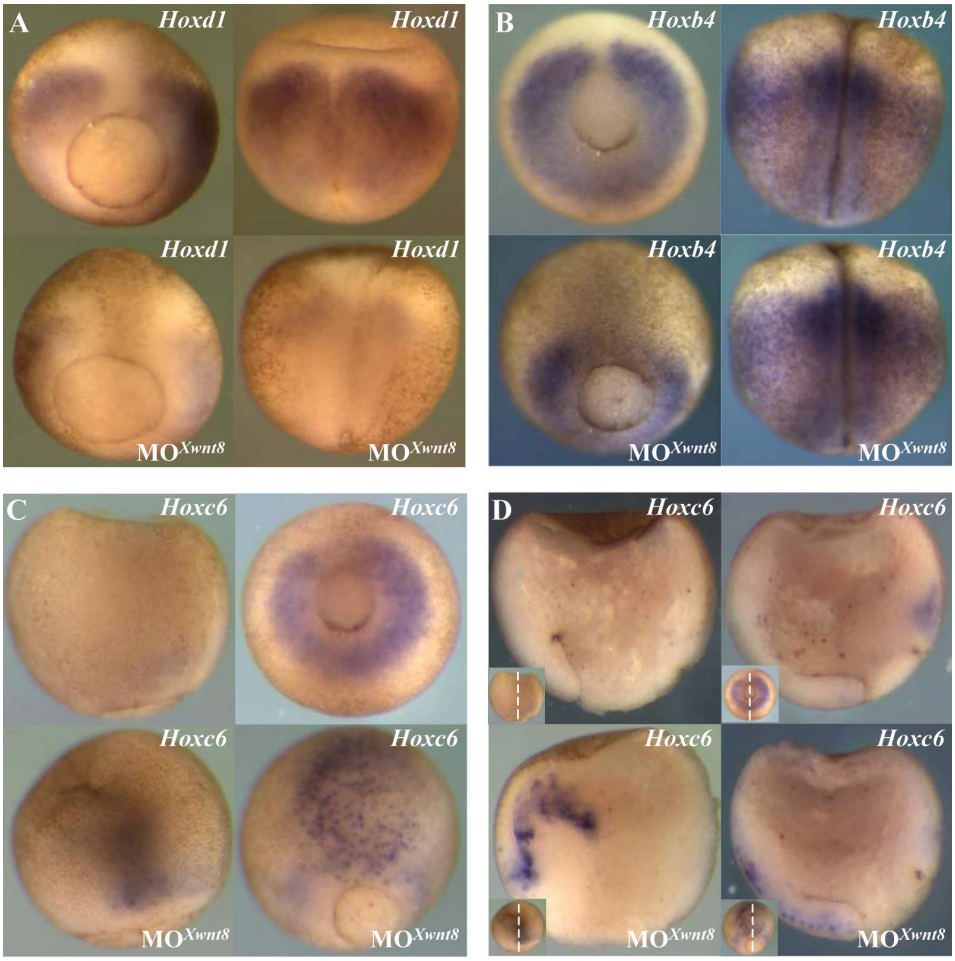

Figure 4. Effects of Xwnt8 loss-of-function on expression of *Hoxd1*, *Hoxb4*, and *Hoxc6*. Embryos were injected into the animal hemisphere at the one-cell stage with 32 ng of MO^{Xwnt8} and analysed by whole-mount *in situ* hybridisation. Injected embryos are shown at the bottom of each panel, control embryos are shown on top. Shown are vegetal views with dorsal to the top. **(A)** Expression of *Hoxd1*, shown are stages 11 (left side of the panel), and stage 12 (right side of the panel). The effect of injected MO^{Xwnt8} is a downregulation in *Hoxd1* expression. **(B)** Expression of *Hoxb4*, shown at stage 11.5 (left side of the panel) and stage 13 (right side of the panel) appears unaffected by Xwnt8 loss-of-function. **(C)** Expression of *Hoxc6* is upregulated by Xwnt8 loss-of-function on the dorsal side of the embryo. Shown are stages 10.5 (left side of the panel) and 11.5 (right side of the panel). **(D)** Dorsal to ventral sections of the embryos shown in **(C)**, the plane of sectioning is depicted by the dotted line in the insets on the bottom left corner.

that expression of *ceh-13*, the labial ortholog of *C. elegans*, depends on Wnt signalling. Strikingly, regulatory elements of *ceh-13* can act as WG response elements in transgenic *Drosophila* embryos. This points, together with our results, to a conserved and ancient mechanism, wherein labial-type *Hox* gene expression is dependent on Wnt signalling. However, to our knowledge, it has never been reported that induction of labial-type *Hox* genes by Wnt signalling can be accomplished when protein synthesis is inhibited and is therefore direct.

Is *Xwnt8* involved in generating a gradient?

The different effects of *Xwnt8* gain-of-function on the expression of different *Hox* paralog groups may contribute to the generation of an early *Hox* pattern. This pattern is initiated in mesoderm, followed by the appearance of the *Hox* code in neurectoderm. A posterior to anterior positive gradient of β–cat/Wnt signalling in neurectoderm has been suggested (Kiecker and Niehrs, 2001). If a gradient of anteroposterior patterning information spreads from the posterior tissues to more anterior tissues we expect the *Hox* genes to be functional downstream of this gradient and, as a consequence, to respond to changes in this gradient. The observed effects on *Hox* expression by *Xwnt8* loss- and gain-of-function make the existence of such a gradient unlikely. In *Xwnt8* loss-of-function, posterior *Hox* genes would be expected to be downregulated, considering that the source of the gradient is thought to be on the posterior side of the embryo. This is in conflict with the observed induced expression of *Hoxc6* in dorsal mesoderm and ectoderm, and an enhanced level of expression in ventrolateral mesoderm of embryos injected with MO^{Xwnt8}. Unaffected expression of *Hoxb4* as seen in *Xwnt8* loss- and gain-of-function is also not in corroboration with a simple model wherein a gradient of Wnt signalling along the anteroposterior axis is used to provide positional information. According to a model like this, a gene expressed more anteriorly would be expressed in more posterior tissues in response to loss-of-function for the gradient, the results of MO^{Xwnt8} experiments are contradictory; they show a strong downregulation of *Hoxd1* expression in embryos with reduced Xwnt8 signalling, and not a posterior shift. Upregulation of *Hoxc6* expression observed in Xwnt8 gain-of-function is much weaker and in a significantly smaller domain than upregulated *Hoxd1* expression, while leaving expression of *Hoxb4* unaffected, these results also contradict a model whereby an anteroposterior gradient of Wnt signalling is used to pattern the primary axis. We propose a model wherein Xwnt8 is involved in initiating a pattern of *Hox* expression in ventrolateral marginal zone mesoderm of the embryo by

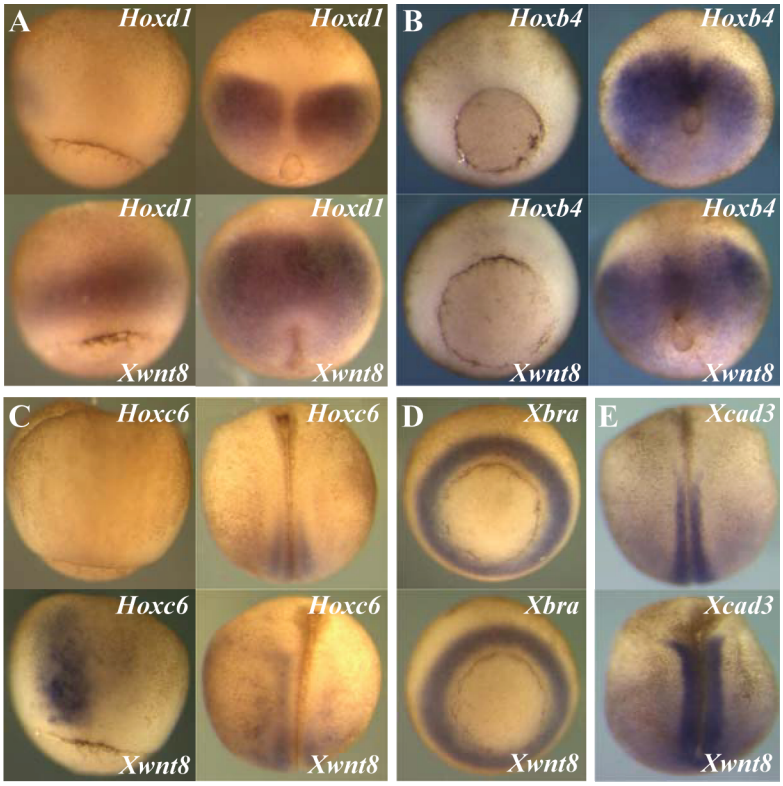

Figure 5. Effects of Xwnt8 gain-of-function on the expression of *Hoxd1*, *Hoxb4*, *Hoxc6*, *Xbra*, and *Xcad3*. Embryos were injected at the one-cell stage into the animal hemisphere with 100 pg CS2-*Xwnt8* plasmid, and analysed by whole-mount *in situ* hybridisation. Injected embryos are shown on the bottom of each panel, control embryos on the top. **(A)** Expression of *Hoxd1* is ectopically upregulated in dorsal tissues of injected embryos. Shown are stage 10 (left side of the panel) and stage 12.5 (right side of the panel) embryos, the views are dorsal with anterior to the top. **(B)** The expression of *Hoxb4* appears unchanged by Xwnt8 gain-of-function, shown are stage 10.5 (left side of the panel,) and stage 12.5 (right side of the panel) embryos, views are vegetal with dorsal to the top. **(C)** Expression of *Hoxc6* is upregulated in dorsally at stage 10 (left side of the panel), and in neurectoderm of stage 15 (right side of the panel) embryos. **(D)** Expression of the mesodermal marker *Xbra* appeared unaltered, shown are stage 11 embryos in vegetal view with dorsal up. **(E)** Expression of the posterior marker *Xcad3* is shifted to a more anterior position as a result of the Xwnt8 gain-of-function, shown are embryos at stage 17, dorsal view, with anterior up.

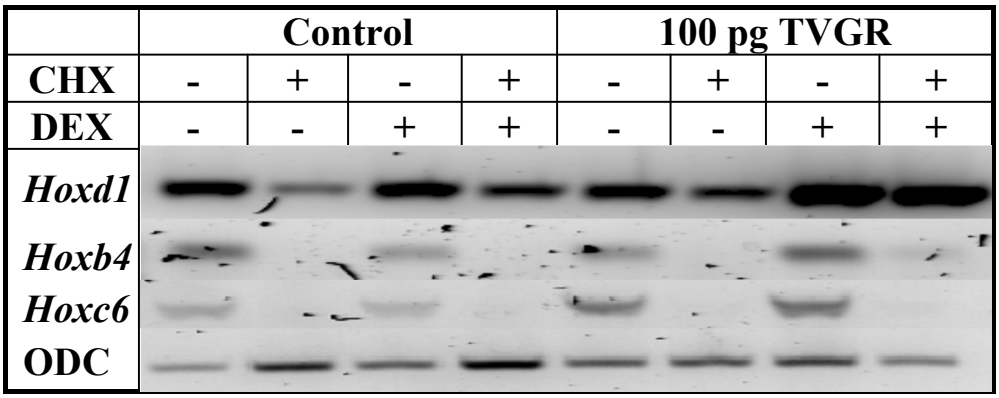

Figure 6. Tcf/Lef signalling is directly upstream of expression of *Hoxd1*. Embryos at the one-cell stage were injected into the animal hemisphere with 100 pg TVGR, an activated hormone inducible form of XTcf3. Cycloheximide (CHX) was added before the start of gastrulation, followed by addition of dexamethasone (DEX), see for details the materials and methods section. In control embryos expression of *Hoxd1*, *Hoxb4*, and *Hoxc6* was repressed or inhibited by addition of CHX; addition of DEX on the other hand did not lead to a difference in expression of the three *Hox* genes assayed, expression in the combined CHX and DEX treatment appears as in the only CHX treatment. Injection of TVGR and subsequent addition of CHX has no effect on expression of the three *Hox* genes assayed. Activation of TVGR by DEX however, led to an induction of *Hoxd1* expression. The induction of *Hoxd1* is shown to be direct by addition of DEX in presence of CHX.

initiating the expression of *Hoxd1*, in a direct fashion. After the initial activation, the Hox cascade continues, *Hoxb4* expression is initiated in a nested domain to *Hoxd1*, in a manner independent of canonical Wnt signalling, thereby creating the first step of the Hox code. This leads to our conclusion that Xwnt8, and perhaps other Wnts, plays an important part in setting up the *Hox* code. In fact this code of colinearly expressed *Hox* genes can be considered as a gradient of positional information along the anteroposterior axis.

Materials and methods

Xenopus embryos and microinjections

Pigmented *Xenopus laevis* embryos were obtained by *in vitro* fertilisation, and after dejelling in a 2% cysteine solution (pH 8.0), cultured in 0.1x Marc's Modified Ringers's (MMR) (Sive *et al.*, 2000), containing 50 µg/ml gentamycin at 14-21 ^{o}C. Embryos were injected in 1x MMR + 4% ficoll and afterwards transferred to 1x MMR + 1% Ficoll, and cultured in this medium for 1 to 7 hours, after which they were transferred and to 0.1x MMR in which they were cultured until harvesting. Staging of the embryos was performed according to Nieuwkoop and Faber (Nieuwkoop and Faber, 1967).
MO^{Xwnt8}, supplied by Gene Tools, LLC, has the sequence: 5'-tttgcatgatgaaggctgctatccg. The MO^{contr} has the sequence: 5'-cctcttacctcagttacaatttata. Embryos were injected into the animal pole at the one-cell stage with 32 or 64 ng MO^{Xwnt8}, dissolved in water, in a volume of 4 or 8 nl respectively, or with the MO^{contr} using the same conditions.
The CS2-*Xwnt8* construct was made by cloning the full-length coding region of *Xwnt8*, obtained by PCR using the CSKA-X8 plasmid (Christian and Moon, 1993) as template and the following primers: f: 5'-gaggaattccggatagca gccttcatcatgcaaaacacc, r: 5'–ctactcgagtctccggtggcctctgttcttcc, containing an *Eco*RI or a *Xho*I restriction site respectively, in the CS2+ vector (Rupp *et al.*, 1994) using the restriction sites in the primers. 50 pg, in a volume of 8 nl, of this plasmid was injected, dissolved in water, into the animal pole of embryos at the one-cell stage.
Synthetic capped mRNA was made using the Ambion mMessage mMachine Kit, 100 pg TVGR mRNA was injected into the animal pole of one-cell stage embryos.
Treatments with Cycloheximide (CHX) and Dexamethasone (DEX) were performed using concentrations of 10 µM. CHX treatment was started 30

minutes prior to DEX addition. Embryos were harvested 4 hours after addition of CHX (Kolm and Sive, 1995a).

Whole mount *in situ* hybridisation and antisense probes

Whole mount *in situ* hybridisations were performed (Harland, 1991), with minor modifications. The antisense RNA probes were generated by run off *in vitro* translation using DIG RNA labelling mix (Roche), and T7 or Sp6 RNA polymerase (Promega). The probes were generated using the following templates: *Hoxd1*: (Sive and Cheng, 1991), *Hoxb4*: a 708 bp fragment containing the complete *Hoxb-4* ORF cloned in pGEMTE, *Hoxc6*: a 998 bp *Hoxc-6* fragment in pGEM1 containing a part of the homeodomain and extending into the 3' UTR, *Xcad3*: (Pownall *et al.*, 1996); *Xbra*: pSP73Xbra (Smith *et al.*, 1991).

RT-PCR

Total RNA was extracted using Tri-Pure reagent (Roche). First strand cDNA was subsequently synthesised using Superscript KSII polymerase (Gibco-BRL), primed with an Oligo dT15 according to the manufacturer's instruction. RT-PCR assays were performed in the exponential phase of amplification as described (Busse and Séquin, 1993) using Tfl polymerase (Promega) in buffer containing 20 mM TrisAc, pH 9.0, 75 mM KAc, 10 mM NH_4SO_4, 1.7 mM $MgSO_4$ and 0.05% Tween-20. The primers used are: *Hoxd1*: f: 5'-agggaactttgcccaactctcc r: 5'-gtgcagtacatgggtgtctggc; *Hoxb4* (Hooiveld *et al.*, 1999); *Hoxc-6* f: 5'-cagagccagacgtggactattcatccagg; Hoxc-6 r: 5'-caaggtaactgtcacagtatggagatgatggc; ODC f: 5'-gtcaatgatgggtgtatggatc r: 5'- tccattccgctctcctgagcac.

Acknowledgments

We thank C. McNulty, G. Mainguy, and S. Wacker for critically reading the manuscript. PIdR was supported by a NWO-grant and EU "HPRN" and "Quality of life" programs.

References

Bailey, W. J., Kim, J., Wagner, G. P., Ruddle, F. H., (1997), Phylogenetic reconstruction of vertebrate Hox cluster duplications. Mol. Biol. Evol. 14: 843-853.

Bouillet, P., Oulad-Abdelghani, M., Ward, S. J., Bronner, S., Chambon, P., Dolle, P., (1996), A new mouse member of the Wnt gene family, mWnt-8, is expressed during early embryogenesis and is ectopically induced by retinoic acid. Mech. Dev. 58: 141-152.

Bürglin T. R. and Ruvkun G., (1993), The *Caenorhabditis elegans* homeobox gene cluster. Curr. Opin. Genet. Dev. 3: 615-620.

Bürglin, T. R., Ruvkun, G., Coulson, A., Hawkins, N. C., McGhee J. D., Schaller, D., Wittmann, C., Müller, F., Waterston, R. H., (1991), Nematode homeobox cluster. Nature 350: 703

Busse, U. and Séquin C., (1993), Molecular analysis of the Wnt-1 proto-oncogene in Ambystoma mexicanum (axolotl) embryos. Differentiation 53: 7-15.

Christian, J. L., McMahon, J. A., McMahon, A. P., Moon, R. T., (1991), *Xwnt-8*, a *Xenopus Wnt-1/int-1* related gene responsive to mesoderm inducing growth factors, may play a role in ventral mesodermal patterning during embryogenesis. Development 111: 1045-1055.

Christian, J. L., Moon, R. T., (1993), Interactions between *Xwnt8* and Spemann organiser signalling pathways generate dorsoventral pattern in the embryonic mesoderm of *Xenopus*. Genes Dev. 7: 13-28.

Darken, R. S. and Wilson, P. A., (2001), Axis induction by Wnt signalling: Target Promoter Responsiveness Regulates Competence. Dev. Biol. 234:42-54.

De Robertis, E. M., Oliver, G., and Wright, C. V. E., (1989), Determination of Axial Polarity in the vertebrate Embryo: Homeodomain Proteins and Homeogenetic Induction. Cell 57: 189-191.

Derrigo, M., Cestelli, A., Savettieri, G., Di Liegro, I., (2000), RNA-protein interactions in the control of stability and localization of messenger RNA (review). Int. J. Mol. 5: 111-23.

Duboule, D., (1994), Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of the morphologies through heterochrony. Development Suppl.: 135-142

Erter, C. E., Wilm, T. P., Basler, N., Wright, C. V. E., Solnica-Krezel, L., (2001), Wnt8 is required in lateral mesodermal precursors for neural posteriorization in vivo. Development 128: 3571-3583.

Fredieu, J. R., Yanzhen, C., Maier, D., Danilchik, V., Christian, J. L., (1997), Xwnt-8 and Lithium Can Act upon either Dorsal Mesodermal or Neurectodermal Cells to Cause a Loss of Forebrain in *Xenopus* Embryos. Dev. Biol. 186: 100-114.

Gaunt S. J. and Strachan, L., (1996), Temporal Colinearity in expression of Anterior Hox Genes in Developing Chick Embryos. Dev. Dyn. 207: 270-280.

Gaunt, S. J., Sharpe, P. T., Duboule, D., (1988), Spatially restricted domains of homeo-gene transcripts in mouse embryos: relation to a segmented body plan. Development Suppl. 104: 169-179.

Glinka, A., Wu, W., Delius, H., Monaghan, A. P., Blumenstock, and Niehrs, C., (1998), Dickkopf-1 is a member of a new family of secreted proteins and functions in head induction. Nature 391: 357-362.

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

Harland, R. M., (1991), *In situ* hybridisations: an improved whole mount method for *Xenopus* embryos. In Methods in Cell Biology, Vol 36 (ed. B. K. Kay and H. J. Peng), 685-695. San Diego, CA: Academic Press.

Harvey, R. P., Melton, D. A., (1988), Microinjection of synthetic Xhox-1A homeobox mRNA disrupts somite formation in developing Xenopus embryos. Cell 1988 53: 687-697.

Heasman, J., (2002), Morpholino Oligos: Making Sense of Antisense? Dev. Biol. 243: 209-214.

Hooiveld, M. H. W., Morgan, R., In der Rieden, P., Houtzager, E., Pannese, M., Damen, K., Boncinelli, E., and Durston, A. J., (1999), Novel interactions between vertebrate *Hox* genes. Int. J. Dev. Biol. 43: 665-674.

Hoppler, S., Brown, J. D., and Moon, R. T., (1996), Expression of a dominant-negative Wnt blocks induction of MyoD in Xenopus embryos. Genes Dev. 10: 2805-2817.

Hume, C. R. and Dodd, J., (1993), *Cwnt-8C*: a novel *Wnt* gene with a potential role in primitive streak formation and hindbrain organization. Development 119: 1147-1160.

Kiecker, C. and Niehrs, C., (2001), A morphogen gradient of Wnt/β-catenin signalling regulates anteroposterior neural patterning in *Xenopus*. Development 128: 4189-4201.

Kolm, P. J. and Sive, H. L., (1995a), Efficient hormone-inducible protein function in *Xenopus laevis*. Dev. Biol. 171:267-272.

Kolm, P. J. and Sive, H. L., (1995b), Regulation of the *Xenopus labial* Homeodomain Genes, HoxA1 and HoxD1: Activation by Retinoids and Peptide Growth Factors. Dev. Dyn. 167: 34-49.

Kühl, M., Finnemann, S., Binder, O., Wedlich, D., (1996), Dominant negative expression of a cytoplasmically deleted mutant of XB/U-cadherin disturbs mesoderm migration during gastrulation in Xenopus laevis. Mech. Dev. 54: 71-82.

Lawrence, P. A. and Morata, (1994), Homeobox genes: their function in Drosophila segmentation and pattern formation. Cell 78: 181-189.

Lekven, A. C., Thorpe, C. J., Waxman, J. S., and Moon, R. T., (2001), Zebrafish *wnt8* Encodes Two Proteins on a Bicistronic Transcript and Is Required for Mesoderm and Neurectoderm Patterning. Developmental Cell 1: 103-114.

Manak, J. K. and Scott, M. P. (1994), A class act: conservation of homeodomain protein functions. Development Suppl., 61-77.

McGinnis, W. and Krumlauf, R., (1992), Homeobox genes and axial patterning. Cell 68: 283-302.

Molenaar, M., Van de Wetering, M., Oosterwegel, M., Peterson-Maduro, J., Godsave, S., Korinek, V., Roose, J., Destree, O., Clevers, H., (1996), XTcf-3 Transcription Factor Mediates β-Catenin-Induced Axis Formation in *Xenopus* Embryos. Cell 86, 391-399.

Nieuwkoop, P. D. and Faber, J., (1967), Normal table of *Xenopus laevis* (Daudin). Amsterdam: North Holland.

Oliver, G., Wright, C. V., Hardwicke, J., De Robertis, E. M., (1988), Differential antero-posterior expression of two proteins encoded by a homeobox gene in Xenopus and mouse embryos. EMBO J. 7: 3199-3209.

Pownall, M. E., Tucker, A. S., Slack, J. M. W., and Isaacs, H. V., (1996), eFGF, Xcad3 and Hox genes form a molecular pathway that establishes the anteroposterior axis in *Xenopus*. Development 122: 3881-3892.

Ruddle F. H., Bentley, M., Murtha, M. T., Risch, N., (1994) Gene loss and gain in the evolution of the vertebrates. Development (Supplement): 155-161.

Rupp, R. A. W., Snider, L., Weitraub, H., (1994), *Xenopus* embryos regulate the nuclear localisation of XMyoD. Genes Dev. 8: 1311-1323.

Salic, A. N., Kristen, L. K., Evans, L. M., Kirchner, M. W., (1997), Sizzled: a secreted antagonist expressed in the ventral marginal zone of *Xenopus* embryos. Development 124, 4739-4748.

Schughart, K., Kappen, C., Ruddle F. H., (1989), Duplication of large genomic regions during the evolution of vertebrate homeobox genes. Proc. Natl. Acad. Sci. U S A 86: 7067-7071.

Sive, H. L., and Cheng, P. F., (1991), Retinoic acid perturbs the expression of *Xhox.lab* genes and alters mesodermal determination in *Xenopus laevis*. Genes Dev. 5: 1321-1332.

Sive, H. L., Grainger, R. M., Harland, R. M., (2000), Early development of *Xenopus laevis* A laboratory manual. Cold Spring Harbor Laboratory press.

Smith, W. C., and Harland, R. M., (1991), Injected Xwnt-8 RNA acts Early in Xenopus Embryos to Promote Formation of a Vegetal Dorsalizing Center. Cell 67: 753-765.

Smith, J. C., Price, B. M., Green, J. B., Weigel, D., and Herrmann, B. G., (1991), Expression of a *Xenopus* homolog of *Brachyury* (T) is an immediate-early response to mesoderm induction. Cell. 67: 79-87.

Sokol, S., Christian, J. L., Moon, R. T., Melton, D.A., (1991), Injected Wnt RNA induces a Complete Body Axis in Xenopus Embryos. Cell 87: 741-752.

Streit, A., Kohler, R., Marty, T., Belfiore, M., Takacs-Vellai, K., Vigano, M. A., Schnabel, R., Affolter, M., and Müller, (2002), Conserved Regulation of the Caenorhabditis elegans labial/Hox1 Gene ceh-13. Dev. Biol. 242: 96-108.

Turner, D. L. and Weintraub, H., (1994), Expression of achaete-scute homolog 3 in Xenopus embryos converts ectodermal cells to a neural fate. Genes Dev. 8: 1434-1447.

Wacker, S. A., Jansen, H. J., McNulty, C. L., Houtzager, E., and Durston, A. J., Timed interactions between a Hox cascade and the Spemann organiser generate positional information during vertebrate gastrulation. Submitted.

Chapter 3

Analysis *in silico* of retinoid response element conservation in vertebrate Hox cluster identifies *Hox* specific binding sites

G. Mainguy, P. M. J. In der Rieden, E. Berezikov,
J. Woltering, R. H. A. Plasterk and A. J. Durston,

Hubrecht Laboratorium,
Nederlands Instituut voor Ontwikkelingsbiologie,
Uppsalalaan 8, 3584 CT, Utrecht, The Netherlands.

Abstract

Retinoic acid (RA) plays an important role in regulating *Hox* gene expression in vertebrates. We performed a comparative analysis *in silico* of the four *Hox* complexes of human and mouse and identified 40 putative RA-response elements (RAREs). Our strategy based on a search for small sequences not only revealed a high degree of conservation (90% between mouse and human) but also led to the identification of sites present near all of the members of a particular paralogy groups, namely groups 3, 4 and 5. Unexpectedly, only a subset of the possible RARE sequences was present in the Hox clusters. Phylogenetic studies revealed that this bias in site usage was presumably already present prior to gnathostome radiation. We analysed the flanking sequences upstream and downstream of ca. 15,000 annotated homolog gene pairs between human and mouse, and found that this subset of RARE sequences is nearly exclusively found in the *Hox* genes regulatory regions. These Hox cluster specific response elements raise the intriguing possibility that transcriptional regulation of different biological functions could be coordinated at the genomic level via a binding site sequence code.

Introduction

Hox genes encode transcription factors known to play a key role in specifying positional identity along the AP axis and other embryonic axes in metazoan embryos. They are expressed in an ordered spatial and temporal sequence correlating with this from 3' to 5' in chromosomal clusters (5,6). *Hox* expression boundaries correspond with structural identity, and modifications of expression domains coincide with changes in morphological structure along the body axes. Understanding the regulatory mechanisms of *Hox* gene expression is therefore of primary interest in the field of developmental and evolutionary biology (7,8).

Retinoic acid (RA) and its derivatives (collectively named retinoids) play an important role in regulating *Hox* gene expression (reviewed in (9-11)). Retinoids are thought mainly to exert their activities at the transcriptional level, acting as ligands to activate two families of transcription factors, the RA receptors (RAR α, β, and γ and the retinoid X receptors RXR α, β, and γ) It is traditionally considered that the formation of heterodimers between RAR and RXR members is induced by high-affinity response elements containing sequences closely related to a "direct repeat" (DR) of AGGTCA, with characteristic inter-half-site-intervals of 5 nucleotides (DR 5) and to a lesser extent 1 or 2 nucleotides (DR1,2) (see (1,10,12) for review but also (13-15)

for other reported mechanisms of transduction). Dimerisation is crucial as half sites have only a very weak affinity for monomers (12,16). Besides the well established importance of DR1,2 and 5 for mediating RA response, there is now also an increasing body of evidence which supports regulation of transcription by retinoids via non DR sites (see table 1 for sequences and references).

To date, studies carried out both *in vitro* and *in vivo* have led to the identification of 16 functional RA-response elements (RARE) at 9 different positions in the flanking regions of 6 mammalian *Hox* genes (reviewed in (2,17) and see figure 1). A conserved sequence in the promoters of *Hoxa5*, *Hoxb5* and *Hoxc5* has also been reported to mediate an indirect RA response (4). Identification of direct RA-response elements mainly in the flanking regions of paralogy groups (PGs) 1 and 4 is in contrast with the colinear RA response from PG 1 to PG 9 observed in mammalian cultured cells ((18-21)), or from the PG 1 to the PG 5 in animal cap assay in *Xenopus* (22). Furthermore, *Hox* gene activation by RA in the absence of protein synthesis has been observed for *Hoxa1* and *Hoxb1* which have identified RAREs but also for *Hoxd1, Hoxb2* and *Hoxb8* which have not (18,20,23-25). Retinoid-responsive *Hox* genes and RARE bearing *Hox* genes thus, form two partially overlapping but distinct sets, clearly revealing how incomplete our understanding of the RA action is.

Comparison of the human and mouse genomes is a promising method for identifying regulatory elements or transcription factor binding sites (26,27). Because they are composed of 12 nucleotides, RAREs are sufficiently complex to be searched for on a genomic scale. However, the heterogeneity of these response elements (see table 1) precludes the design of a reliable matrix. To identify RAREs in the Hox clusters with minimal assumption, we set up a three steps strategy defined by progressive expansion based on conservation analysis. We first searched the Hox complexes for RARE sequences that were known in the context of other promoters to be functional. This led to identification of sequences with the highest degree of confidence (first-rated sequences). Second, putative modified sites were searched for by local alignment to identified RAREs present at a particular position in only one of the two species. This method allowed the identification of RARE types not described previously (second-rated sequences). Finally, all the sequences present in the Hox clusters (first- and second-rated) were used to establish restricted consensus sequences. This led to the identification of other putative RAREs from which sequences at positions conserved between human and mouse were retained (third-rated sequences).

Results

RAREs are extremely conserved in Hox clusters

Searching in both the human and mouse Hox clusters for the 56 characterised RARE types listed in table 1 led to the identification of 53 putative sites including, indeed, the 20 RA response elements already described (2,4,17) (figure 1). The 33 newly identified sites are very likely to be of physiological interest both for statistical and for phylogenetic reasons. First, this number of sites is much higher than statistically expected by chance (namely 0.04 per type for four clusters, ca. 600,000 bp). Second, most of the sites we found are located at similar positions with respect to particular Hox ORFs in human and mouse, defining 21 conserved positions. Interestingly, this proportion even reaches 100% in the case of the experimentally defined sequences that were identified previously.

Since our first approach identified only exactly matching sequences, natural variants could have been missed. Assuming that function is transitive through conservation, we searched for putative homologues of the 11 unilaterally represented RAREs (figure 1). By performing BLAST searches of surrounding sequences, we identified 3 new types of putative RAREs (in blue on figure 1). Local alignments also revealed that a type 18 DR2 site in the mouse has a type 17 homolog in human in an element 3' to *Hoxa1* (figure 2), as 70% of the nucleotides are identical in the 200 bp surrounding the RARE.

Finally, we used the DR5 RARE sequences identified in the Hox clusters to design a consensus sequence (table 2) and thereby found 11 more putative RAREs. We restricted our analysis to the sites that were conserved between human and mouse and thereby, accepted three new types of putative RAREs (type N1 to N3, table 2). (The same analysis for DR2 and DR1 sequences did not bring any additional conserved sites, not shown). These sites define three new conserved positions in the HoxA clusters, two of which are present at comparable locations in the HoxA cluster of the shark *Heterodontus francisci* (28,29) (figure 2). In summary, we identified 60 RAREs of which 91% are conserved between human and mouse and which define 27 conserved positions within the Hox clusters, and we found six putative RARE sequences which are very likely to be functional RAREs.

Type	Categories	Species	Sequences	References
0	DR5	Sy	AGGTCA(n5)AGGTCA	R03941
1	MGP	Hs	GGTTCA(n5)TGTTCA	(1)
2	Shh	Br	GGTTCA(n5)GGGTCA	(46)
3	**Hoxa1, Hoxb1**	**Hs, Mm**	**GGTTCA(n5)AGTTCA**	**(47)**
4	**Hoxd4 3'**	**Hs, Mm**	**GGTTCA(n5)AGGACA**	**(17)**
5	ADH3	Hs	GGGTCA(n5)AGTTCA	R03945,R03017
6	RARβ2 distal	Hs, Mm	GGGTCA(n5)AGGTTA	(48)
7	RARγ2	Hs	GGGTCA(n5)AGGTGA	R03929
8	CRABPII	Hs	GGGTCA(n5)AGGACA	(49)
9	RARα2	Hs, Mm	AGTTCA(n5)AGTTCA	R04075
10	**Hoxb4 3'**	**Mm**	**AGTTCA(n5)AGGCCA**	**(50)**
11	**Hoxa4, Hoxd4**	**Hs, Mm**	**AGGTGA(n5)AGGTCA**	**(51,52)**
11a	Hoxb4 5'	Mm	GGGTGA(n5)AGGTCA	This work
12	CMV-IEP	Hs	AGGTCA(n5)TGGGCA	(1)
13	CP-H	Mm	AGGTCA(n5)AGGGCA	(1)
14	Ga1	Hs	AGGGCA(n5)AGGTCA	(1)
15	hApoAI	Hs	GGGTCA(n2)GGTTCA	R00135
16	Erythropoietin	Mm	GGGTCA(n2)AGGTCA	(16)
17	CRABPII	Mm	GAGTCA(n2)AGGTCA	R04770
18	CRABPII	Mm	AGTTCA(n2)AGGTCA	R04769
19	CRBPI	Mm	AGGTCA(n2)AGGTCA	R03932
20	CRBPI	Rn	AGGTCA(n2)AAGTCA	R03962
21	**Hoxb1 3'**	**Hs, Mm**	**AGGTAA(n2)AGGTCA**	**(53)**
22	**Hoxb1 5'**	**Hs, Mm**	**AGGGCA(n2)AGTTCA**	**(54)**
23	PAFR 2	Hs	AGGCCA(n2)AAGCCA	(55)
24	ApoCIII	Hs	TGGGCAnAGGTCA	(56)
25	Ovalbumin	Gg	GTGTCAnAGGTCA	(1)
26	HBV	Hs	GGGTAAnGGTTCA	R03934
27	Oct-3/4	Mm	GGGCCAnAGGTCA	(57)
28	PEP	Rn	CGGCCAnAGGTCA	R01690
29	Osteocalcin	Hs	AGGTGAnTCACCG	R01187
30	αB-crystallin	Mm	AGGTCAnGGGTTT	(58)
30a	HoxA4 3'	Hs	AGGTCAnGGGATT	This work
31	MHCI	Mm	AGGTCAnGGGTGG	(1)
32	CRBPII	Mm	AGGTCAnAGGTCA	R04768
33	CRBPII	Mm	AGGTCAnAGTTCA	R03931
34	Stromelysin	Hs	AGGTCAnAGGTCA	(59)
35	ApoAI	Hs	AGGGCAnGGGTCA	(1)
36	CRABPII	Mm	AGGGCAnAGGTCA	(1)
37	Acyl-CoA	Rn	AGGACAnAGGTCA	R03986

38	Oct-3/4	Mm	AGGTCAAGGCTA	(57)
39	TrePal	Sy	AGGTCATGACCT	(41)
40	Growth Hormone	Bt	GGGACATGACCC	(1)
41	MCAD	Hs	GGGTAAAGGTGA	(1)
42	RBP	Hs	CGGTGANTCAGG	(60)
43	Cdx1	Mm	AAGGGTCGTGACCCCT	(61)
44	Vitellogenin A2	Xl	AGGTCA(n3)TGACCT	(1)
45	Laminin B1	Mm	AGGTGA(n3)AGGTTA	(62)
46	αB-crystallin	Mm	GTGTCA(n3)TGCCAA	(58)
47	Cardiac α-	Hs,Rn	AGGTGA(n4)AGGACA	R03942,R01169
48	RBP	Hs	AGGACA(n4)CGGGCA	(60)
49	Laminin B1	Mm	GGGTCA(n5)TCTGGC	(62)
50	αB-crystallin	Mm	GTGTCA(n5)CCAAAT	(58)
51	HSV-1 tk	Hs	AGGTGA(n6)TGGCCT	R04759
52	IR7	Hs, Br	RGGTCA(n7)TGACCY	(63)
53	γF-crystallin	Mm	TGACCC(n8)AGGTCA	R03963
54	Laminin B1	Mm	GGTTAA(n13)GGGTCA	(62)
I	**Hox5 BS**	**Hs, Mm**	**RKCACGTGAYTC**	**(4)**
Ia	HoxA4 3'	Mm	GTCATGTGACTC	This work

Table 1. RARE sequences set. The different sections display functional RAREs (DR5, DR2, DR1 and non DR sites) identified in regulatory elements in *Hox* genes (**in bold**) and non-*Hox* genes respectively. Each of these sites presents a different sequence. Complex sites were also included but, when possible, were narrowed to a core sequence of 12 nucleotides in order to be able to search for reasonable complexity. References or accession number in TRANSFAC database are given. Sequence polymorphisms identified in this study are shaded (see below). Note that type I is not a RXR-RAR binding site and is likely to be an indirect RA response element (4). (Br: *Brachydanio rerio*; Bt: *Bos taurus*; Gg: *Gallus gallus*; Hs: *Homo sapiens*; Mm: *Mus musculus*; Rn: *Rattus norvegicus*; Sy: *Synthetic;* Xl: *Xenopus laevis*).

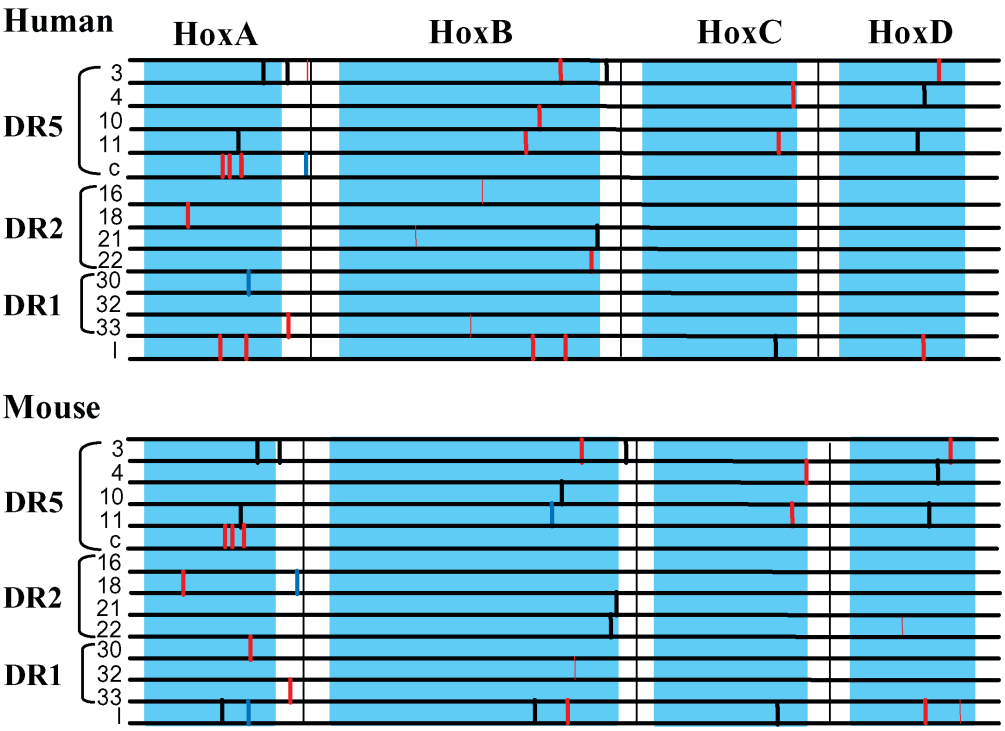

Figure 1. First-rated RAREs in Hox clusters are strongly conserved between human and mouse. Graphic overview of identified RARE locations in human and mouse clusters. The four clusters are defined by the stop and the ATG codons of the first and last paralogs respectively, and theirs positions are represented by light blue-shaded boxes. They are oriented 5' to 3' such as the first PGs are on the right. We included 20 kb on both sides of each clusters (in white, separated by dashed lines). Sites are divided in DR5, DR2 and unconventional RAREs. Previously known and new RAREs are drawn in black and red respectively. Conserved elements from human to mouse are in bold. Numbers refer to RARE type sequences (see table 1).

Distribution of RAREs is asymmetric along the Hox clusters

Several RAREs in the flanking regions of *Hoxa1*, *Hoxb1*, *Hoxa4*, *Hoxb4,* and *Hoxd4* have been shown to mediate a RA response for these genes (depicted in black on figure 1). Here we identify new putative RAREs around PG 1 and PG 4 emphasising the importance of retinoid regulation for these PGs. Moreover, the highly conserved RARE sequence 5' to *Hoxb4* is, at the same time, present in all of the vertebrate *Hox* clusters studied, and is found only upstream of PG 4, thus defining a signature for this PG (see figure 3A). The new conserved RARE identified 5' to PG 3, is a surprising result as, to our knowledge, there is so far no functional evidence that any paralog of this group is directly activated by RA. Moreover this RARE is found in all 3 paralogs studied and is extremely conserved not only at the level of the site itself but also of the surrounding regions (see figure 3B). However, even though these RAREs lie within the PG 3 proximal regions, they could act as long range enhancers as they are located in a region shown to influence the transcription of PG 4 members (17).

A striking feature is that 54 of the 60 response elements (both direct and indirect) are distributed between the first and the fifth PGs (see figure 2). Moreover sites 5' to PG 5 are not conserved between human and mouse except for a DR2 located within the intron of *Hoxa10* which is also present in the shark. This asymmetric partition of RAREs along the *Hox* clusters suggests that high sensitivity to retinoids is somehow restricted to the five anterior PGs. This statement is in extraordinary accordance with experimental findings in *Xenopus* that *Hoxb1* through *Hoxb5* are colinearly induced by RA in early embryogenesis (22). This finding together with our discovery of a high degree of RARE conservation between the mammalian and shark HoxA clusters (figure 2) suggests that RA regulation of anterior *Hox* genes is an ancient and conserved feature of vertebrates.

Interestingly, an asymmetry is also observed in the repartition of the different DR5 types along the clusters. In fact, all of the DR5 types found in the Hox clusters can be grouped in two categories depending on their first half site sequences (table 2). It is then remarkable that the subclass A members (RGTTCA) are spread out from PG 1 to PG 4 while subclass B members (RGGTGA) are restricted between PG 4 and PG 5 defining two abutting domains within the Hox clusters.

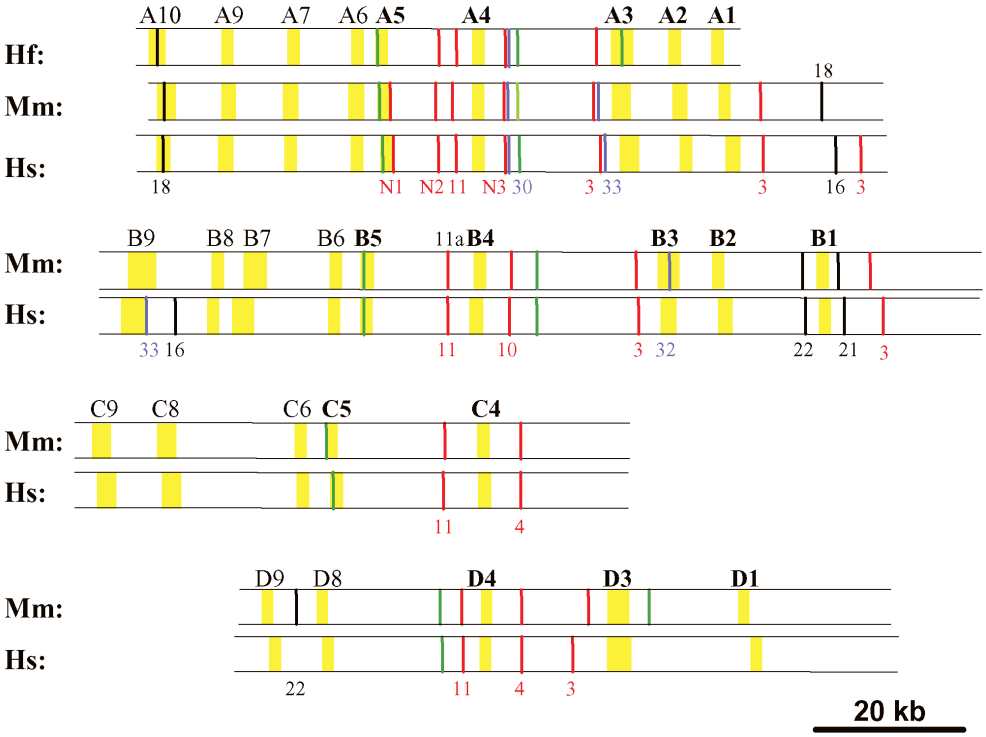

Figure 2. RAREs are mainly distributed in the region between PG 1 and PG 5. For each cluster, sand-coloured boxes represent genes of the PGs 1 to 9. Posterior genes are not represented as no RARE has been found beyond the PG 10. RAREs are drawn by vertical slashes and associated numbers refer to the site-sequence described in table 2. DR5 and DR2 are in red and black respectively. Indirect response elements (Hoxc5 BS, I) are in green (Mm: *Mus musculus*; Hs: *Homo sapiens*; Hf: *Heterodontus francisci*).

A conserved bias in RARE sequence usage in the Hox clusters

It is not only the positions of particular RAREs relative to transcription units that are conserved, but as seen from figure 3, the sequences of the RAREs themselves are also strongly constrained. All direct response elements found belong to the DR classes and 66% of these sequences are DR5. No palindromic, inverted, everted or other atypical sequences from our database matched even though 17 of those sequences were used in the analysis (see table 1). As all the sites have the same complexity this result indicates a preference, and could reflect a specificity of action. Surprisingly, sequences are also biased within the DR5 class. In fact, the first screen resulted in the finding that all of the 25 identified DR5s present a sequence previously first described in *Hox* genes (as illustrated by their names in table 1), while 11 equally natural sequences are absent. A trivial explanation could be that binding site-sequences are similar due to their common origin and their subsequent conservation. However, this seems unlikely as variations in the sequence of homologous binding sites are observed (see figures 2 and 3), and identical types are found at different positions in the same cluster (e.g. type 3 present at three different conserved positions in the HoxA cluster). In contrast, despite the smaller subset of natural sequences available in the DR2 class, we found RAREs first characterised in non-*Hox* genes, namely in the erythropoietin and CRABPII promoters (type 16 and 18). These sites are likely to be functional as they define two conserved positions and are interchangeable in the vicinity of *Hoxa1*.

Therefore, it seems that not only Hox clusters present a specific subset of DR5 RAREs, but also reciprocally these sequences are, at least to some extent, Hox cluster specific. If they were not, RAREs present in *Hox* genes should also have been discovered in non-*Hox* genes. If this is indeed true, these RAREs should not be found in non-*Hox* gene promoters. To test this challenging hypothesis, we undertook a comparative analysis of conserved sites in the promoters of murine and human genes. As a proof of principle we searched for conserved sites within flanking regions of 15,000 annotated gene pairs between human and mouse, including *Hox* genes of the A, B and D clusters. As RARE polymorphism is observed at several positions in the Hox clusters, we searched for the presence of identical or different Hox DR5 types (listed in table 2). Using these criteria we identified 15 positive pairs of conserved RAREs. *Hox* genes (*Hoxa1*, *Hoxa4*, *Hoxb4*, *Hoxd4*, *Hoxa3*, *Hoxb3* and *Hoxd3*) represent 66% of the positive pairs and the five non-Hox genes encode meplin, Poly(RC)-binding protein, CGI-112, γ-crystallin, and oestrogen receptor γ. The identification of only five non-Hox positive pairs of

Site name	Sequence	Occurrence
DR5	AGGTCA(n5)AGGTCA	
Type 3	GGTTCA(n5)AGTTCA	Mm: 5 Hs: 6
Type 4	GGTTCA(n5)AGGACA	Mm: 1 Hs: 1
Type 10	AGTTCA(n5)AGGCCA	Mm: 1 Hs: 1
Type 11	AGGTGA(n5)AGGTCA	Mm: 3 Hs: 4
Type 11a	GGGTGA(n5)AGGTCA	Mm: 1 Hs: 0
Consensus	**RGKTSA(n5)AGKHCA**	Mm: 18 Hs: 16
Type N1	GGGTGA(n5)AGGCCA	Mm: 1 Hs: 1
Type N2	GGTTCA(n5)AGTCCA	Mm: 1 Hs: 1
Type N3	AGTTCA(n5)AGGACA	Mm: 1 Hs: 1

Table 2. DR5 RARE types present in the clusters. First lines (in black) display first and second-rate sequences from which consensus sequences are derived. New conserved sequences found with the DR5 consensus follow (third-rate, type N1 to 3, in red). Exceptional nucleotides are in red and invariable nucleotides are in blue.

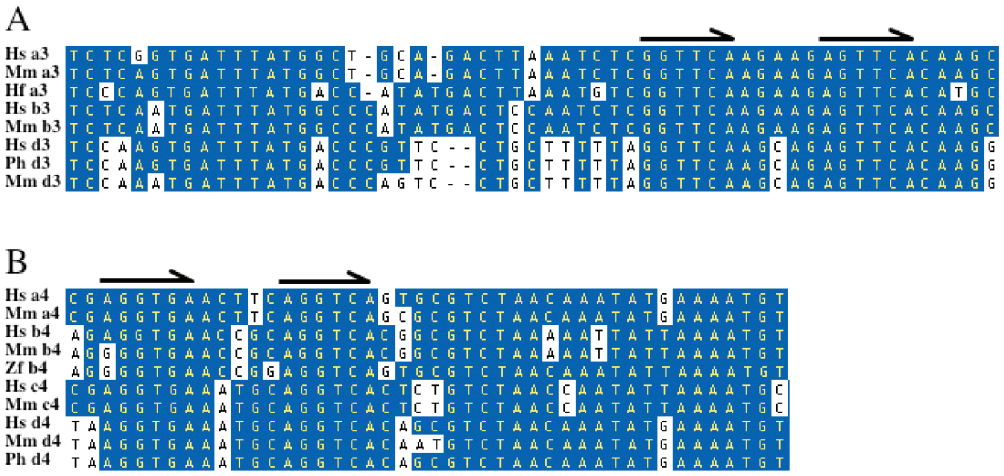

Figure 3. Conserved region associated with RARE of the third (A) and the fourth (B) PGs. Half sites are represented by arrows. Br: *Brachydanio rerio;* Hs: *Homo sapiens*; Mm: *Mus musculus;* Ph: *Papilio hamadryas;* Hf: *Heterodontus francisci*.

genes out of the 15,000 analysed, confirms the bias in site usage and demonstrates that these site sequences are almost *Hox* gene specific.
As the regulation of *Hox* genes by RA seems to be an ancient characteristic of the vertebrates ((30) and cf. supra), it was engaging to test whether this is also the case for the RARE usage. In the nematode and the *Tribolium* clusters we could not find any putative RARE whatsoever, a consistent finding as ecdysozoans are non likely to possess any true retinoid receptors (31,32). Interestingly however, the Drosophila ANTP-C and BX-C clusters display five DR5 derived consensus RAREs even though none of our 60 primary sequences are present. In the shark HoxA cluster, the 4 DR5 RAREs found are of *Hox*-types and no RARE belongs to a non-*Hox* type (figure 2). As the ancestors of *H. francisci* and human diverged at least 400 Myr ago, this strongly suggests that this peculiar site usage was already present prior to gnathostomes radiation. This ancient origin is in agreement with the fact that in mammals, the four clusters are all biased, even though they originate before the divergence of agnathans and gnathosotomes (33).

Discussion

High conservation of RAREs

The objectives of this work were double. We wanted to make an inventory of RAREs of known types in the Hox complexes of human and mouse and to design a method to find new types. Our strategy based on a search for small sequences not only revealed an unexpected high degree of conservation of RAREs between different vertebrates but also led to the identification of particular RAREs which were located close to all members of the same PG, namely PGs 3, 4 and 5. Due to their relatively small size, these conserved motifs cannot be easily detected by extensive alignment. For instance, aligning HoxA and HoxD clusters using PipMaker (http://bio.cse.psu.edu/pipmaker, not shown), failed to identify the conserved RAREs illustrated in figure 3 even though this degree of conservation is comparable to the homeobox encoding regions. Therefore our method is suitable for identifying small functional sequences and proves to be a useful supplement to alignment for identification of phylogenetic footprints (29), as exemplified by our finding of 8 more RAREs conserved between the shark and mammalian HoxA clusters (28).
One important finding from our study is the unexpected extent to which RA-response elements are conserved between different Hox clusters in the same species. In fact, despite the similarities in their expression patterns,

differences between *Hoxb4* and *Hoxa4* in the positions and activities of their regulatory regions had previously led to the proposal that control elements might not be conserved between these two paralogs (34-36). More generally, based on comparative studies of HoxA clusters in vertebrates it has been suggested that cis-acting elements are rarely conserved after duplication (29). It had nonetheless been reported that *Hoxa4*, *Hoxd4* and *Hoxb4* have maintained the position of a 3' neural enhancer (37). Our findings here now demonstrate extensive conservation of regulatory elements between all of the Hox clusters. Therefore, conservation of these sequences despite post-duplication drift strongly suggests that they are essential for *Hox* gene functions, which are ancestral for gnathostomes. This level of conservation is all the more striking because the endogenous expression of *Hoxb4* in the mouse can virtually be recreated without its 5' RARE element (34). Along these lines, the indirect response element upstream of PG 5 is present in the *Hoxd5* gene of *H. francisci* (not shown), and is also present in the HoxD clusters of both human and mouse even though *Hoxd5* has been lost in those species, suggesting a function of the RARE which is independent of this particular gene. One exciting possibility could be that these elements are important for the biology of the complex itself, namely its colinearity.

Nature of the bias usage of retinoid binding sites in Hox complexes

We report here that only a few of the possible RXR-RAR binding sites are actually distributed throughout the clusters and reciprocally that these sequences are essentially unavailable outside Hox genes. Pressure for maintaining such specific sequences could be due to interaction with co-factors or to specific heterodimers.
Indeed, the promoter context of a given RARE is an important factor in eliciting its transcriptional response, as exemplified by the fact that despite the *Hoxa1* and *RARβ2* genes bearing an identical RARE, their responses to retinoids can be dissociated (38). In this context, it is noteworthy that in our case, nucleotide conservation is consistently extended beyond the two canonical half sites, as 93% (73/80) of the inter-space nucleotides are identical between human and mouse for the 15 pairs of conserved DR5. Moreover, in the case of the RARE 5' to PG 3, the conserved sequence is 23 base pairs long with only 3 observed substitutions (figure 3A), and in some cases, the inter-space nucleotide sequence seems to be more critical than the half-sites (figure 3B). The specific site usage could therefore indicate the presence of RXR-RAR interacting co-factors (coincident pressure), or that the site might be bound by additional transcription factors (serial pressure).

The RAR and RXR families each contain three different genes each with several isoforms (reviewed in (39,40)), and despite an important network of compensation and redundancy between RXR-RAR heterodimers, there are arguments for differential activity of certain heterodimers on specific target genes (41-44). For example, in the murine F9 cell line, inductions of *Hoxa1* and *Hoxb1* by RA are differentially mediated as only *Hoxa1* induction depends on RARγ (42). Furthermore, different RAREs require different heterodimers for optimal activation at a given ligand concentration (41). Therefore, the special usage of binding sites could reveal a recognition code, namely certain sequences could preferentially be bound by certain RXR-RAR heterodimers containing specific isoforms. The fact that a given heterodimer can differentially regulate different RAREs depending on the type of retinoid or on the ligand concentrations (41), raises the intriguing possibility that the ligand could also be a source of variation used as a recognition code, even though to our knowledge there is no experimental evidence for such a fine-tuned mechanism within the DR5 class of response elements, as discrimination of sequences via specific heterodimer-ligand complexes has been reported for different structured RAREs (41). In this view, Hox specific sequences allow detection, via *bona fide* receptors, of particular retinoids or specific concentrations of a given ligand. This latter hypothesis is of special interest since this mechanism would account for the high sensitivity of *Hox* gene inductions to retinoid concentrations (reviewed in (21)) and would give a framework to our findings that two classes of regulatory elements divide the clusters into two sections. The previous hypotheses are not mutually exclusive and restriction of the RARE sequences could reflect action of a particular combination of receptors, co-receptors and ligands, different parameters that could be regulated independently.

Retinoids regulate very different biological processes in a broad range of tissues during development and adulthood and ultimately the same genome responds differently each time. The specificity of the response is though to be achieved via different combinations of DNA-interacting proteins, but the existence of a predictable DNA-protein recognition code is a controversial issue (45). Our findings that some binding site sequences are associated with a function suggests that understanding how specific functions are discriminated as different blueprints on the DNA may not be out of reach.

Methods

Genomic DNA sources

The genomic sequences of the Hox clusters were retrieved from the Ensembl databases (http://www.ensemble.org) maintained by the Wellcome Trust Sanger Institute and the European Bioinformatics Institute. Our analysis were performed using the following sequences : human HoxC and HoxD, v.6.28.1 and HoxA and HoxB v.7.29a.3; mouse HoxA and HoxD,v.6.3a.1 and HoxB and HoxC v.7.3b.3. This combination of clusters results in 0.01% and 0.28% of ambiguous sequences for human and mouse respectively. GenBank accession numbers for the other clusters: *Heterodontus francisci* HoxM (HoxA-like) cluster (AF479755; AF224262); *Caenorhabditis elegans,* (CEY75B8A; CEY79H2); *Tribolium castaneum* partial sequence of the ANTP-C (AY043292; AY043293; AF321227); *Drosophila melanogaster* BX-C (DMU31961) and ANTP-C (AE001572); *Papio hamadryas* HoxB cluster (AC116664).

RAREs data set

Experimentally defined RAREs were assembled from previous collections by (1,2) and from the transcription factor database TRANSFAC 5.0 (3) (http://transfac.gbf.de/TRANSFAC/cl/cl.html), in association with a literature survey (Table 1). Each sequence is associated with a type number and referred as such in the text. Each included sequence has been demonstrated to respond to RA *in vivo* either in transgenic animals or in a cell-culture model system. Synthetic sequences for which no biological effect has been demonstrated were not taken into account. We also included a sequence known to mediate indirect RA response on *Hoxc5* and to be conserved in *Hoxa5*, *Hoxb5* and *Hoxc5* (4) (Hox5 binding site: Hox5 BS).

RAREs identification procedures

Searches for specific RARE types listed in table 1 were performed with MapDraw (DNAstar, Inc). For sites present only in the mouse or human clusters, putative homologs were systematically searched using 200 bp of the surrounding region of unilateral RAREs as probes by BLAST search and local alignment using Clustral method (Megalign; DNAstar, Inc). Sequences of RAREs effectively present in the complex allowed the design of consensus sequences for each DR classes that were used in a second round of analysis (MapDraw).

High throughput promoter analysis.

Considering the distribution of RAREs in respect to gene positions in the Hox clusters, search was performed within 5 kb upstream and 5 kb downstream of the gene start sites. A list of human and mouse orthologous genes has been obtained from ENSEMBL server and 10 kb flanking sequences (5 upstream, 5 downstream) of ca. 15,000 orthologs have been extracted. Search for RARE types listed in table 2 have been implemented via simple PERL script, and promoter pairs bearing one type each (even different) have been reported. Accession numbers for positive non-Hox genes are given for human, RAREs being upstream (meplin ENSG00000050405, Poly(RC)-binding protein ENSG00000090097) and downstream (CGI-112 ENSG00000100908; γ-crystallin ENSG00000127377; and oestrogen-related receptor γ, ENSG00000057103).

Acknowledgements

This study was supported in part by HPRN and EU quality of life programs and G.M. is a recipient of an EMBO postdoctoral fellowship.

References

1. Mangelsdorf, D.J., Umesono, K. and Evans, R.M. (1994) In Sporn, M. B., Roberts, A. B. and Goodman, D. S. (eds.), The Retinoids: Biology, Chemistry and Medicine. Second Edition ed. Raven Press, New York, pp. 319-349.
2. Huang, D., Chen, S.W. and Gudas, L. J. (2002) Dev. Dyn., 223, 353-370.
3. Wingender, E., Chen, X., Hehl, R., Karas, H., Liebich, I., Matys, V., Meinhardt, T., Pruss, M., Reuter, I. and Schacherer, F. (2000) Nucleic Acids Res., 28, 316-319.
4. Arcioni, L., Simeone, A., Guazzi, S., Zappavigna, V., Boncinelli, E. and Mavilio, F. (1992) EMBO J.. 11, 265-277.
5. Duboule, D. and Morata, G. (1994) Trends Genet., 10, 358-364.
6. Krumlauf, R. (1994) Cell, 78, 191-201.
7. Duboule, D. (1994) Dev. Suppl, 135-142.
8. Gaunt, S.J. (2000) Int. J. Dev. Biol., 44, 109-117.
9. Hofmann, C. and Eichele, G. (1994) In Sporn, M. B., Roberts, A. B. and Goodman, D. S. (eds.), The Retinoids: Biology, Chemistry and Medicine. Second Edition ed. Raven Press, New York, pp. 387-441.
10. Durston, A.J., van der Wees, J., Pijnappel, W.W. and Godsave, S.F. (1998) Curr Top Dev. Biol., 40, 111-175.
11. Gavalas, A. and Krumlauf, R. (2000) Curr. Opin. Genet. Dev., 10, 380-386.
12. Rastinejad, F. (2001) Curr. Opin. Struct. Biol., 11, 33-38.
13. Willy, P. J., Umesono, K., Ong, E.S., Evans, R. M., Heyman, R. A. and Mangelsdorf, D. J. (1995) Genes Dev., 9, 1033-1045.
14. Perlmann, T. and Jansson, L. (1995) Genes Dev., 9, 769-782.
15. Kobayashi, M., Yu, R.T., Yasuda, K. and Umesono, K. (2000) Mol. Cell. Biol., 20, 8731-8739.
16. Makita, T., Hernandez-Hoyos, G., Chen, T. H., Wu, H., Rothenberg, E. V. and Sucov, H. M. (2001) Genes Dev., 15, 889-901.
17. Zhang, F., Nagy Kovacs, E. and Featherstone, M. S. (2000) Mech. Dev., 96, 79-89.
18. Simeone, A., Acampora, D., Arcioni, L., Andrews, P. W., Boncinelli, E. and Mavilio, F. (1990) Nature, 346, 763-766.
19. Papalopulu, N., Lovell-Badge, R. and Krumlauf, R. (1991) Nucleic Acids Res., 19, 5497-5506.

20. Simeone, A., Acampora, D., Nigro, V., Faiella, A., D'Esposito, M., Stornaiuolo, A., Mavilio, F. and Boncinelli, E. (1991) Mech. Dev., 33, 215-227.
21. Boncinelli, E., Simeone, A., Acampora, D. and Mavilio, F. (1991) Trends Genet., 7, 329-334.
22. Godsave, S. F., Koster, C. H., Getahun, A., Mathu, M., Hooiveld, M., van der Wees, J., Hendriks, J. and Durston, A.J. (1998) Dev. Dyn., 213, 39-49.
23. LaRosa, G.J. and Gudas, L.J. (1988) Mol. Cell. Biol., 8, 3906-3917.
24. Kolm, P.J. and Sive, H. (1995) Dev. Biol., 167, 34-49.
25. Lu, H. C., Revelli, J. P., Goering, L., Thaller, C. and Eichele, G. (1997) Development, 124, 1643-1651.
26. Wasserman, W. W., Palumbo, M., Thompson, W., Fickett, J. W. and Lawrence, C. E. (2000) Nat. Genet., 26, 225-228.
27. Loots, G. G., Ovcharenko, I., Pachter, L., Dubchak, I. and Rubin, E. M. (2002) Genome Res., 12, 832-839.
28. Kim, C. B., Amemiya, C., Bailey, W., Kawasaki, K., Mezey, J., Miller, W., Minoshima, S., Shimizu, N., Wagner, G. and Ruddle, F. (2000) Proc. Natl. Acad. Sci. U. S. A., 97, 1655-1660.
29. Chiu, C. H., Amemiya, C., Dewar, K., Kim, C. B., Ruddle, F. H. and Wagner, G. P. (2002) Proc. Natl. Acad. Sci. U. S. A., 99, 5492-5497.
30. Manzanares, M., Wada, H., Itasaki, N., Trainor, P. A., Krumlauf, R. and Holland, P. W. (2000) Nature, 408, 854-857.
31. Escriva, H., Safi, R., Hanni, C., Langlois, M. C., Saumitou-Laprade, P., Stehelin, D., Capron, A., Pierce, R. and Laudet, V. (1997) Proc. Natl. Acad. Sci. U. S. A., 94, 6803-6808.
32. Sluder, A. E., Mathews, S. W., Hough, D., Yin, V. P. and Maina, C. V. (1999) Genome Res., 9, 103-120.
33. Force, A., Amores, A. and Postlethwait, J. H. (2002) J. Exp. Zool., 294, 30-46.
34. Whiting, J., Marshall, H., Cook, M., Krumlauf, R., Rigby, P. W., Stott, D. and Allemann, R. K. (1991) Genes Dev., 5, 2048-2059.
35. Behringer, R. R., Crotty, D. A., Tennyson, V. M., Brinster, R. L., Palmiter, R. D. and Wolgemuth, D. J. (1993) Development, 117, 823-833.
36. Gould, A., Morrison, A., Sproat, G., White, R. A. and Krumlauf, R. (1997) Genes Dev., 11, 900-913.
37. Morrison, A., Ariza-McNaughton, L., Gould, A., Featherstone, M. and Krumlauf, R. (1997) Development, 124, 3135-3146.

38. Taneja, R., Roy, B., Plassat, J. L., Zusi, C. F., Ostrowski, J., Reczek, P. R. and Chambon, P. (1996) Proc. Natl. Acad. Sci. U. S. A., 93, 6197-6202.
39. Mollard, R., Viville, S., Ward, S. J., Decimo, D., Chambon, P. and Dolle, P. (2000) Mech. Dev., 94, 223-232.
40. Parrado, A., Despouy, G., Kraiba, R., Le Pogam, C., Dupas, S., Choquette, M., Robledo, M., Larghero, J., Bui, H., Le Gall, I. et al. (2001) Nucleic Acids Res., 29, 4901-4908.
41. La Vista-Picard, N., Hobbs, P. D., Pfahl, M. and Dawson, M. I. (1996) Mol. Cell. Biol., 16, 4137-4146.
42. Chiba, H., Clifford, J., Metzger, D. and Chambon, P. (1997a) Mol. Cell. Biol., 17, 3013-3020.
43. Chiba, H., Clifford, J., Metzger, D. and Chambon, P. (1997b) J Cell Biol, 139, 735-747.
44. Abu-Abed, S. S., Beckett, B. R., Chiba, H., Chithalen, J. V., Jones, G., Metzger, D., Chambon, P. and Petkovich, M. (1998) J. Biol. Chem., 273, 2409-2415.
45. Benos, P. V., Lapedes, A. S. and Stormo, G. D. (2002) Bioessays, 24, 466-475.
46. Chang, B. E., Blader, P., Fischer, N., Ingham, P. W. and Strahle, U. (1997) EMBO J.. 16, 3955-3964.
47. Langston, A. W., Thompson, J. R. and Gudas, L. J. (1997) J. Biol. Chem., 272, 2167-2175.
48. Valcarcel, R., Holz, H., Jimenez, C.G., Barettino, D. and Stunnenberg, H. G. (1994) Genes Dev., 8, 3068-3079.
49. Astrom, A., Pettersson, U., Chambon, P. and Voorhees, J. J. (1994) J. Biol. Chem., 269, 22334-22339.
50. Gould, A., Itasaki, N. and Krumlauf, R. (1998) Neuron, 21, 39-51.
51. Popperl, H. and Featherstone, M. S. (1993) Mol. Cell. Biol., 13, 257-265.
52. Mavilio, F. (1993) Eur. J. Biochem, 212, 273-288.
53. Marshall, H., Studer, M., Pöpperl, H., Aparicio, S., Kuroiwa, A., Brenner, S. and Krumlauf, R. (1994) Nature, 370, 567-571.
54. Studer, M., Pöpperl, H., Marshall, H., Kuroiwa, A. and Krumlauf, R. (1994) Science, 265, 1728-1732.
55. Mutoh, H., Fukuda, T., Kitamaoto, T., Masushige, S., Sasaki, H., Shimizu, T. and Kato, S. (1996) Proc. Natl. Acad. Sci. U. S. A., 93, 774-779.
56. Reue, K., Leff, T. and Breslow, J. L. (1988) J. Biol. Chem., 263, 6857-6864.

57. Pikarsky, E., Sharir, H., Ben-Shushan, E. and Bergman, Y. (1994) Mol. Cell. Biol., 14, 1026-1038.
58. Gopal-Srivastava, R., Cvekl, A. and Piatigorsky, J. (1998) J. Biol. Chem., 273, 17954-17961.
59. Anglard, P., Melot, T., Guerin, E., Thomas, G. and Basset, P. (1995) J. Biol. Chem., 270, 20337-20344.
60. Panariello, L., Quadro, L., Trematerra, S. and Colantuoni, V. (1996) J. Biol. Chem., 271, 25524-25532.
61. Houle, M., Prinos, P., Iulianella, A., Bouchard, N. and Lohnes, D. (2000) Mol. Cell. Biol., 20, 6579-6586.
62. Sharif, K. A., Li, C. and Gudas, L. J. (2001) Mech. Dev., 103, 13-25.
63. Le Jossic, C. and Michel, D. (1998) Biochem. Biophys. Res. Commun., 245, 64-69.

Chapter 4

Identifying HOX paralog groups by the PBX-binding region

Morgan, R., In der Rieden, P. M. J., Hooiveld, M. H. W., and Durston, A. J.

Hubrecht Laboratorium,

Nederlands Instituut voor Ontwikkelingsbiologie,

Uppsalalaan 8, 3584 CT, Utrecht, The Netherlands.

Reprinted from Trends Genet. (2000) 16, 66-67

Identifying HOX paralog groups by the PBX-binding region

Richard Morgan
morgan@ucl.ac.uk

Paul In der Rieden
rieden@niob.knaw.nl

Michiel H.W. Hooiveld
hooiveld@niob.knaw.nl

Antony J. Durston
tony@niob.knaw.nl

The Hubrecht Laboratory, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands.

The *Hox* genes are a family of transcription factors that define specific anteroposterior identities, both in vertebrate and in invertebrate embryos[1], and that are characterized by a very highly conserved DNA-binding motif known as the homeodomain[2–4]. *In vitro*, most HOX proteins recognize the same four-base-pair consensus sequence that is actually repeated many times in the genome[5]. Far greater binding specificity is achieved when HOX proteins bind as heterodimers with PBX proteins (vertebrate homologs of *Drosophila* homeodomain-containing transcription factor *extradenticle*)[6]. PBX and HOX proteins interact at a specific and highly conserved hexapeptide on the surface of the HOX protein[6–8]. This short sequence of amino acids is necessary for PBX binding and, apart from the homeodomain itself, is the most characteristic feature of *Hox* genes.

During the evolution of vertebrates, the ancestral cluster of *Hox* genes was duplicated at least twice[1,9–11]; hence, most vertebrates have at least four independent *Hox* clusters, referred to as a, b, c…etc. Despite some *Hox* genes in each cluster having become non-functional or even entirely deleted subsequent to the duplication step, the overall genomic structure of each cluster has been conserved in evolution[1]. In general, the descendants of each of the genes in the ancestral *Hox* cluster have similar expression patterns and some conserved functions[12,13]. They are described as paralogs (e.g. *hoxa1*, *hoxb1* and *hoxd1*). Outside of the homeodomain region, the overall sequence identity between members of each paralog group is very low. As a result, the paralog identity of each gene has often been ambiguous.

Two previous studies have addressed this problem by aligning *Hox* genes based on their hexapeptide sequences rather than their homeodomain[14,15] (which forms the usual basis for *Hox* gene alignment comparisons). Their findings revealed that there were, indeed, some amino acids adjacent to the hexapeptide that are conserved only within individual paralog groups. Here we have extended these studies to include all hexapeptide-containing paralog groups from a wide range of species. Interestingly, this reveals that there are several very highly conserved amino acids clustered around the hexapeptide sequence. These amino acids consistently identify *Hox* genes as belonging to a particular paralog group (Fig. 1).

Why are the amino acids around the hexapeptide sequence so highly conserved between paralogs but not clusters? Members of one paralog bind to a distinct DNA sequence only when bound to PBX at the hexapeptide site;

FIGURE 1. Conserved amino acids around the hexapeptide sequence of *Hox* paralog groups (1–8)

```
HOX1      TFDWMKVKRNxPK
HOX2     PEfPWMKEKKxxxK
HOX3    KQIFPWMKESxQxxKxK
HOX4     VVYPWMKKhHhxxxxxxY
HOX5    PQIYPWMrKLHhxHxxxxxxKR
HOX6      hYPWMQRMNSxxxxxfGxxxRxR
HOX7     RIYPWMRSxGxD
HOX8      hfPWM
```

trends in Genetics

Capital letters denote the conserved amino acids (based on the standard IUB codes). The amino acids in bold are those that define the hexapeptide itself. 'f' denotes either phenylalanine (F) or tyrosine (Y) and 'r' denotes either arginine (R) or lysine (K). 'h' is a hydrophobic amino acid. 'x' indicates that the amino acid at that position is not conserved (although the spacing is). The sequences compared for each paralog group were as follows. HOX1: *Hoxa1* (R, H, M, X, Z), *Hoxb1* (F, H, M, X, Z) and *Hoxd1* (M, X). HOX2: *Proboscipedia* (D), *Hoxa2* (M) and *Hoxb2* (H). HOX3: *Hoxa3* (M, Z), *Hoxb3* (CP, C, H, M, X) and *Hoxd3* (C, H, M). HOX4: *Hoxa4* (C, H, M), *Hoxb4* (F, M, X, Z), *Hoxc4* (H, M, Z) and *Hoxd4* (C, H, M, Z). HOX5: *Sex combs reduced* (D), *Hoxa5* (H, M, Z), *Hoxb5* (C, H, M, X, Z) and *Hoxc5* (H, M). HOX6: *Hoxb6* (H, M, Z) and *Hoxc6* (F, H, M, X, Z). HOX7: *Hoxa7* (H, M, Z) and *Hoxb7* (H, M, X). HOX8: *Hoxb8* (C, M, X) and *Hoxc8* (F, H, M, X). Abbreviated species names: D, *Drosophila*; C, chicken; CP, carp; F, *Fugu*; H, human; M, mouse; R, rat; X, *Xenopus*; Z, Zebrafish.

 PII: S0168-9525(99)01881-8

however, this hexapeptide is highly conserved. Thus, paralog-specific amino acids surrounding this motif might 'fine tune' the PBX–HOX binding interaction, resulting in a unique DNA-binding specificity for each paralog group. Indeed, in the case of the *Hox4* paralog group at least, there is already some evidence that this might be the case[15]. In a recent report[16], a *Drosophila* PBX homolog is shown to determine DNA-binding specificity directly via a 21 bp element. This is distinct from the sequence that is bound by the HOX homeodomain. Hence, the function of PBX might not be limited to simply modifying the binding specificity of HOX proteins. We hope that future studies on the requirement for these amino acids close to the hexapeptide might provide clues as to how this 'fine-tuning' works.

References

1 Carroll, S.B. (1995) Homeotic genes and the evolution of arthropods and chordates. *Nature* 376, 479–485
2 McGinnis, W. *et al.* (1984) A conserved DNA sequence in homoeotic genes of the *Drosophila* Antennapedia and bithorax complexes. *Nature* 308, 428–433
3 Scott, M.P. and Weiner, A.J. (1984) Structural relationships among genes that control development: sequence homology between the *Antennapedia*, *Ultrabithorax*, and *fushi tarazu* loci of *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* 81, 4115–4119
4 Scott, M.P. *et al.* (1989) The structure and function of the homeodomain. *Biochim. Biophys. Acta Rev. Cancer* 989, 25–48
5 Hayashi, S. and Scott, M.P. (1990) What determines the specificity of action of *Drosophila* homeodomain proteins? *Cell* 63, 883–894
6 Chang, C-P. *et al.* (1996) Pbx modulation of Hox homeodomain amino-terminal arms establishes different DNA-binding specificities across the *Hox* locus. *Mol. Cell. Biol.* 16, 1734–1745
7 Piper, D.E. *et al.* (1999) Structure of a HoxB1–Pbx1 heterodimer bound to DNA: role of the hexapeptide and a fourth homeodomain helix in complex formation. *Cell* 96, 587–597
8 Passner, J.M. *et al.* (1999) Structure of a DNA-bound Ultrabithorax–Extradenticle homeodomain complex. *Nature* 397, 714–719
9 Kappen, C. *et al.* (1989) Two steps in the evolution of *Antennapedia*-class vertebrate homeobox genes. *Proc. Natl. Acad. Sci. U. S. A.* 86, 5459–5463
10 Scott, M.P. (1993) A rational nomenclature for vertebrate homeobox (HOX) genes. *Nucleic Acids Res.* 21, 1687–1688
11 Meyer, A. and Malaga-Trillo, E. (1999) Vertebrate genomics: More fishy tales about Hox genes. *Curr. Biol.* 9, 210–213
12 Horan, G.S. *et al.* (1995) Mutations in paralogous Hox genes result in overlapping homeotic transformations of the axial skeleton: evidence for unique and redundant function. *Dev. Biol.* 169, 359–372
13 Condie, B.G. and Capecchi, M.R. (1994) Mice with targeted disruptions in the paralogous genes hoxa-3 and hoxd-3 reveal synergistic interactions. *Nature* 370, 304–307
14 Sharkey, M. *et al.* (1997) Hox genes in evolution: protein surfaces and paralog groups. *Trends Genet.* 13, 145–151
15 Shanmugam, K. *et al.* (1997) Residues flanking the HOX YPWM motif contribute to cooperative interactions with PBX. *J. Biol. Chem.* 272, 19081–19087
16 Li, X. *et al.* (1999) A sequence motif distinct from Hox-binding sites controls the specificty of a Hox response element. *Development* 126, 5581–5589

Chapter 5

XMeis3 is necessary for mesodermal *Hox* gene expression

Paul M. J. In der Rieden and Antony J. Durston

Hubrecht Laboratorium,
Nederlands Instituut voor Ontwikkelingsbiologie,
Uppsalalaan 8, 3584 CT, Utrecht, The Netherlands.

Abstract

Hox transcription factors provide positional information during patterning of the anteroposterior axis. Recently, an early expression sequence of *Hox* genes was found in *Xenopus* gastrulae mesoderm. Hox transcription factors can co-operatively bind with PBC-class co-factors, enhancing specificity and affinity for consensus binding sites. The nuclear localisation of these co-factors is regulated by the Meis-class of homeodomain proteins. During development of the zebrafish hindbrain, *Meis3* has been shown to synergise with *Hoxb1* in the autoregulation of *Hoxb1*. In *Xenopus XMeis3* posteriorises the embryo upon ectopic expression. We investigated whether *XMeis3* is involved in regulation of *Hox* gene expression in mesoderm during gastrulation. Here, we present evidence that *XMeis3* is necessary for expression of *Hoxd1*, *Hoxb4* and *Hoxc6* in mesoderm during gastrulation. In addition, we show that *XMeis3* function is necessary for the progression of gastrulation. Finally, we propose synergy between *XMeis3* and *Hoxd1* in *Hoxd1* autoregulation during gastrulation.

Introduction

During the development of most animal species studied, *Hox* transcription factors specify positional information along the anterior to posterior axis (Bürglin *et al.*, 1991; Bürglin and Ruvkun, 1993; McGinnis and Krumlauf, 1992; Lawrence and Morata, 1994; Manak and Scott, 1994). *Hox* genes form a subfamily of the homeobox containing gene family, and are organised in four clusters, each located on different chromosomes. The homeobox encodes a DNA binding motif called the homeodomain. A strict control of the expression and function of these *Hox* genes is essential. It has been shown that Pbx family members, and their *Drosophila melanogaster* counterpart *Extradenticle* (Exd), function as cofactors for Hox proteins; they can enhance their binding specificity and affinity for target sequences on DNA (van Dijk and Murre, 1994; Knoepfler and Kamps, 1995; Chang *et al.*, 1995; Neuteboom and Murre, 1997; Ryoo and Mann, 1999). Pbx/Exd family members are part of a different subfamily of the homeodomain containing proteins, namely the TALE-class. This class is characterised by having a three amino acid loop extension between the first and second helices of their homeodomains (Bürglin, 1997). Cooperative binding of Hox and Pbx/Exd proteins can lead to transactivation while binding of the individual factors leads to repression on the same promoter elements (Pinsonneault *et al.*, 1997). When Hox proteins bind to DNA cooperatively with a Pbx/Exd family

member, the main protein-protein interaction consists of binding of the hexapeptide motif of the Hox protein to a pocket formed by the atypical homeodomain of PBC family members (Jabet *et al.*, 1999; Passner *et al.*, 1999; Piper *et al.*, 1999). This pocket is composed of the three amino acid loop extension of the PBC homeodomain, residues in the third helix of the homeodomain, and a residue in the C-terminal helix of PBC homeodomains (Piper *et al.*, 1999). The nuclear localisation of Pbx/Exd proteins is controlled by competing nuclear import and export signals (Abu-Saar *et al.*, 1999). When members of the Meis family, or their *Drosophila* counterpart *Homothorax* (Hth), also members of the TALE-class of homeodomain proteins, are present in the cytoplasm they interact with Pbx/Exd family members in such a way that the nuclear export signal of the Pbx/Exd family member is shielded, resulting in a net influx of Pbx/Exd into the nucleus, influencing the function of Hox proteins present (Ryoo *et al.*, 1999; Ryoo and Mann, 1999; Jaw *et al.*, 2000). However, Pbx/Exd and Meis/Hth proteins are not used exclusively as cofactors for Hox proteins. The myogenic bHLH factors (Knoepfler *et al.*, 1999) and Engrailed (Peltenburg and Murre, 1996) also depend on the activity of Pbx and Meis members for proper functioning.

For Hox paralog group 1 members, autoregulation dependent on Pbx/Exd and Meis/Hth has been shown to function in neurectoderm of mouse embryos (Pöpperl *et al.*, 1995; Ferretti *et al.*, 2000), *C. elegans* (Streit *et al.*, 2002) and in endoderm of *Drosophila* embryos (Ryoo *et al.*, 1999; Marty *et al.*, 2001). Binding of Hox and Pbx family members to bipartite Hox-Pbx binding sites is essential for autoregulation (Pöpperl *et al.*, 1995; Grieder *et al.*, 1997; Ryoo *et al.*, 1999; Marty *et al.*, 2001). Meis proteins have been shown to be indispensable as mediators of this process (Grieder *et al.*, 1997; Ryoo *et al.*, 1999; Marty *et al.*, 2001).

In *Xenopus*, a member of the Meis family, *XMeis3*, is a posteriorising factor in neurectoderm of *Xenopus laevis*, and is required for hindbrain patterning (Salzberg *et al.*, 1999; Dibner *et al.*, 2001). In zebrafish embryos, similar functions have been reported for Meis3 and other Meis family members (Vlachakis *et al.*, 2001; Waskiewicz *et al.*, 2001; Choe *et al.*, 2002). Expression of *XMeis3* is reported as being initiated in a stripe in the neural plate of early-mid neurula embryos. During neurula and early-tailbud stages, expression is mainly localised to rhombomeres (r's) 2, 3, and 4, and the anterior spinal cord, while posterior rhombomeres show some ventral expression (Salzberg *et al.*, 1999). Expression of *XMeis3* overlaps with neurectodermal expression of *Hoxd1* (r4 and r5) (Kolm and Sive, 1995), *Hoxb4* (r7, r8, and the anterior spinal cord) (Harvey and Melton, 1988), and *Hoxc6* (anterior spinal cord) (Oliver *et al.*, 1988; De Robertis *et al.*, 1989).

These overlaps are consistent with the idea that *XMeis3* is involved in controlling the function of Hox proteins it is co-expressed with. These studies do, however, leave many questions unanswered. They pay little attention to when and where Meis cofactors actually interact with Hox proteins at different stages during the early AP patterning process. These details are likely to be crucial for understanding the mechanism at hand. Studies of vertebrate *Hox* expression and function have already made it clear that AP patterning depends on a specific early spatiotemporal sequence of *Hox* gene expression. Expression of each Hox gene is initiated in a specific domain in the gastrula embryo and then undergoes an establishment phase during which this expression domain changes to a gene specific AP zone in axial mesoderm and the neural plate and finally a maintenance phase during which this AP zone is consolidated. This sequence is employed universally in mammals, birds, fish and amphibians and shows generic features in these different species (Duboule and Dolle, 1989; Graham *et al.*, 1989; Gaunt and Strachan, 1996; Deschamps *et al.*, 1999; Wacker *et al.*, submitted). A recent study analysed the early *Hox* expression patterns in *Xenopus*, and this revealed a spatiotemporally colinear initiation of expression of a sequence of *Hox* genes within a horseshoe-shaped domain of ventrolateral marginal zone mesoderm at different stages during gastrulation and then sequential dorsalisation of each *Hox* expression zone corresponding with its translation into a stable AP zone in axial mesoderm and the neural plate (Wacker *et al.*, submitted). This sequence reflects timed interactions between a ventrolateral mesodermal *Hox* cascade and the Spemann organiser that are imperative for AP axis formation.

We set out to investigate whether expression of early *Hox* genes depends on the activity of XMeis3 and whether XMeis3 is involved in regulation of expression of these *Hox* genes in mesoderm during gastrulation. In order for *XMeis3* to be able to regulate Hox expression in mesoderm they need to be co-expressed there. We performed whole mount *in situ* hybridisation to study the detailed early expression of *XMeis3* and compared it to the early expression patterns of *Hoxd1*, *Hoxb4*, and *Hoxc6* and found significant co-expression in lateral regions of marginal zone mesoderm, early during gastrulation. To gain further insight into the early function of *XMeis3*, we followed a gain- and a loss-of-function strategy. In the gain-of-function strategy synthetic *XMeis3* mRNA was injected and expression of *Hox* genes was studied. These experiments showed that ectopic expression of *XMeis3* during gastrulation is capable of inducing expression of the assayed *Hox* genes in mesoderm as well as in ectoderm. In the loss-of-function strategy we made use of an antisense morpholino oligonucleotide (reviewed in Heasman,

2002 and references therein) to inhibit the translation of *XMeis3* mRNA (MO^{XMeis3}). Injection of MO^{XMeis3} leads to a reduction in expression of *Hoxd1*, *Hoxb4*, and *Hoxc6* in mesoderm and ectoderm during gastrulation, and to severe patterning defects. Finally we show synergy between Hoxd1 and XMeis3 and show that the mesodermal expression of *Hoxd1* during early gastrulation is already dependent on XMeis3 mediated autoregulation.

Results

The expression of *XMeis3* overlaps with *Hox* gene expression in mesoderm

To determine whether *XMeis3* is co-expressed with *Hox* genes in the mesoderm of gastrula embryos, whole mount *in situ* hybridisations were performed for *XMeis3*, *Hoxd1*, *Hoxb4*, and *Hoxc6*. Expression of *XMeis3* is initiated in a horseshoe-shaped domain in ventrolateral marginal zone mesoderm, before stage 11 expression is lost in the ventralmost tissue, resulting in two lateral expression domains on either side of the organiser in mesoderm of early gastrula stage embryos (Fig. 1A). Expression later in gastrulation becomes localised to mesoderm lateral to the midline and overlying ectoderm (Fig. 1B). Early expression of *Hoxd1*, *Hoxb4*, and *Hoxc6* is initiated in ventrolateral mesoderm and each of these genes follows a similar spatiotemporal expression sequence (Wacker *et al.*, submitted). During early phases of gastrulation mesodermal expression of *Hoxd1* (Fig. 1C), *Hoxb4* (Fig. 1E), and *Hoxc6* (Fig. 1G) overlaps with expression of *XMeis3* in the dorsolateral domains of these *Hox* genes (compare Fig. 1A to 1C, 1E, and 1G). At the end of gastrulation the overlap between expression of *Hoxd1* (Fig. 1D) and *XMeis3* (Fig. 1B) in mesoderm is maintained, and newly initiated expression of both genes in the neurectoderm also overlaps. At the same time, the more posteriorly expressed *Hoxb4* (Fig. 1F) and *Hoxc6* (Fig. 1H) only partially overlap *XMeis3* expression (Fig. 1B) in involuted mesoderm. *Hoxb4* expression partially overlaps expression of *XMeis3* in overlying ectoderm (compare Fig. 1F to 1B). These results show that there is indeed an overlap in expression of *XMeis3* and of early *Hox* genes in mesoderm during gastrulation, and that expression of *XMeis3* and *Hoxd1* also overlaps in ectoderm.

XMeis3 gain-of-function upregulates *Hox* gene expression in mesoderm and ectoderm

To investigate whether *XMeis3* is capable of contributing to the regulation of *Hox* gene expression, 2 ng of synthetic mRNA containing the full-length coding region of *XMeis3* was injected into the animal pole of embryos at the one-cell stage. The amount of 2 ng was chosen because this was shown to lead to posteriorisation of injected embryos (Salzberg *et al.*, 1999). The effects on expression of *Hoxd1*, *Hoxb4*, *Hoxc6*, *Xbra*, and the posterior marker *Xcad3* in gastrula stages were assayed by *in situ* hybridisation. The ectopic expression of *Hoxd1* (Fig. 2A) in injected embryos is remarkable because it is found in the region harbouring the Spemann organiser, tissue that normally does not expresses *Hox* genes. The horseshoe-shaped domain of expression is also expanded and expression levels appear to be enhanced. Furthermore expression can be found in ectoderm of the animal cap and mesoderm underlying it, in the form of a streak of expression in contact with the expanded ring of expression around the blastopore (Fig 2A). *Hoxb4* also shows ectopic expression in animal cap ectoderm and expansion of the endogenous expression domain (Fig. 2B), but no closure of the dorsal expression gap neither in organiser mesoderm nor in overlying ectoderm can be observed. Interestingly, induced expression of *Hoxc6* can already be found in dorsal mesoderm at stage 10.25 (Fig. 2C), significantly earlier than endogenous initiation of expression (st11) and like ectopic *Hoxd1* expression, this occurs in dorsal mesoderm. In later stages an expansion of the endogenous horseshoe-shaped expression domain is also found (data not shown). Expression of the mesodermal marker *Xbra* appears unaltered in injected embryos (Fig. 2D), suggesting that changes in *Hox* expression domains are not due to changes in induction of mesoderm, but rather to its patterning. The previously described posteriorising effect of *XMeis3* on neurectoderm is confirmed by anterior expansion of expression of the posterior marker *Xcad3* (Fig. 2E).

XMeis3 loss-of-function downregulates expression of *Hox* genes

To determine whether *XMeis3* function is necessary for initiation and/or establishment of *Hoxd1*, *Hoxb4*, and *Hoxc6* expression, an antisense morpholino oligonucleotide directed against *XMeis3* mRNA (MO^{XMeis3}) was injected into the animal hemisphere of embryos at the one-cell stage. *XMeis3* loss-of-function leads to a loss of trunk structures and defects in axis specification, in a concentration dependent manner. When 12 ng MO^{XMeis3}

was injected a loss of trunk structures and defects in head development and tail formation can be observed, while the anteriormost structure, the cement gland, remains present (Fig. 3B). When 24 ng MO^{XMeis3} was injected, an enlargement of the cement gland was visible accompanied by a stronger loss of trunk structures (Fig. 3C) In half the injected embryos *spina bifida's* are observed, suggesting that the embryos suffer from gastrulation problems. When 32 ng or more MO^{XMeis3} were injected, the embryos arrested during gastrulation at stage 11 (Fig. 3D). Embryos injected with this high dose of MO^{XMeis3} appear unaffected and posses normal looking blastopores until the moment of arrest. This is unlike what would be expected if the arrest was caused by toxicity of an injected agent, this would generally generate a much larger spread in stages at which embryos die or arrest, accompanied by irregular formation of the blastopore. Removal of the vitelline membrane revealed that cells have lost cell-cell contact, but appear round and intact. This suggests that the observed effect is the result of a strong knockdown of *XMeis3* function and not an aspecific effect of MO^{XMeis3}. Injection of the same amount of a control morpholino (MO^{contr}), in sequence unrelated to MO^{XMeis3}, has no outward effects on embryos (data not shown). These findings support the idea that the gastrulation arrest phenotype is a true result of XMeis3 loss-of-function and that XMeis3 is required for patterning (a part of) the primary axis in *Xenopus* embryos.

To further test the specificity of the MO^{XMeis3}, 125 pg of synthetic *XMeis3* mRNA, lacking most of the sequence that the MO^{XMeis3} is complementary to, was co-injected with 32 ng MO^{XMeis3} into the animal hemisphere of embryos at the one-cell stage (Fig. 3F). The exogenous *XMeis3* was able to largely rescue the MO^{XMeis3} phenotype (compare Fig. 3D to 3E, and 3F). In a small number of the co-injected embryos a full recovery of the axis can be observed, sometimes accompanied by a secondary axial outgrowth out of the primary axis, containing somites (Fig. 3G).

The effect of XMeis3 loss-of-function on *Hox* expression was studied by injecting 16 ng MO^{XMeis3} into the animal hemisphere of embryos at the one-cell stage followed by *in situ* hybridisation. To be able to analyse marker expression in late gastrula stage embryos the arrest in gastrulation, observed after injection of a high amount of MO^{XMeis3}, was avoided, by the injection of 16 ng. The XMeis3 loss-of-function leads to downregulation of expression of *Hoxd1* (Fig. 4A), *Hoxb4* (Fig. 4B), and *Hoxc6* (Fig. 4C), in mesoderm and ectoderm. This led to our conclusion that XMeis3 is necessary for *Hox* gene expression in marginal zone mesoderm, and neural plate ectoderm.

Synergy between Hoxd1 and XMeis3

Autoregulation is known to occur among labial type *Hox* genes in hindbrain ectoderm (Pöpperl *et al.*, 1995; Struder *et al.*, 1998), endoderm of *Drosophila* embryos (Grieder *et al.*, 1997; Ryoo *et al.*, 1999), and *C. elegans* (Streit *et al.*, 2002). For a number of these cases it has been shown that this autoregulation is dependent on a Pbx/Hox bipartite binding site in their promoters (Pöpperl *et al.*, 1995; Grieder *et al.*, 1997; Ryoo *et al.*, 1999; Streit *et al.*, 2002). Because nuclear localisation of Pbx family members is dependent on the action of Meis family members and because XMeis3 loss-of-function led to a significant downregulation of *Hoxd1* expression in mesoderm and ectoderm, we suspected that *XMeis3* might be involved in Hoxd1 autoregulation. To test our idea that XMeis3 may mediate autoregulation of labial type *Hox* genes in *Xenopus* development, we co-injected relatively small amounts of synthetic mRNA for *XMeis3* and *Hoxd1* and also injected them separately using double the amount of mRNA. Small amounts of mRNA were used to be able to observe compound phenotypes in co-injected embryos. If a strong effect in embryos injected with only a single messenger was generated this would not have been possible. The embryos injected with only a single synthetic messenger show little or no phenotypic effects, while co-injected embryos show a significant retardation in head development (Fig. 5). This points towards a synergistic relation between Hoxd1 and XMeis3.
To further test this synergy, and to test whether XMeis3-mediated Hoxd1 autoregulation is involved in the establishment of *Hoxd1* expression, we wished to investigate the necessity of Hoxd1 for maintaining *Hoxd1* expression in mesoderm. If XMeis3 activity is needed in early gastrula mesoderm to enhance or alter the function of Hoxd1, then Hoxd1 loss-of-function should generate the same effect on *Hoxd1* expression as XMeis3 loss-of-function. To test whether this is the case, 32 ng MO^{Hoxd1} (McNulty *et al.*, manuscript in preparation) was injected into the equatorial region of the 2 blastomeres making up the presumptive left side of 4-cell stage embryos. The other half of the embryos served as an internal control. This results in a downregulation of expression of *Hoxd1* in mesoderm on the injected side (Fig. 6A). To further test whether establishment of expression of *Hoxd1* needs both Hoxd1 and XMeis3, sub optimal amounts of morpholinos against both messengers were co-injected and injected separately. Embryos were harvested at stage 11 and assayed for *Hoxd1* expression (Fig 6B). Sub optimal morpholino amounts were used to allow different levels of reduction

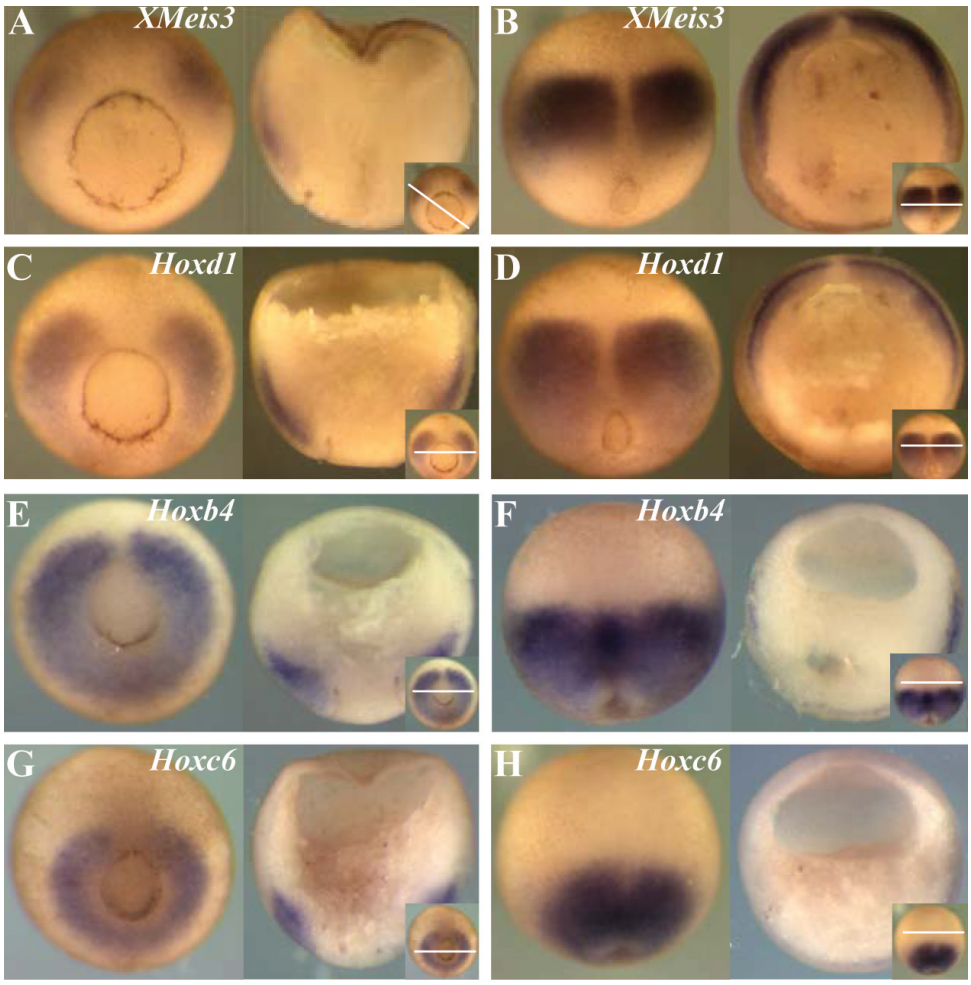

Figure 1. Expression of *XMeis3*, *Hoxd1*, *Hoxb4*, and *Hoxc6* during gastrulation. Embryos were analysed by whole-mount *in situ* hybridisation for expression of *XMeis3* **(A** and **B)**, *Hoxd1* **(C** and **D)**, *Hoxb4* **(E** and **F)**, and *Hoxc6* **(G** and **H)**. Whole mounts are shown on the left side of each panel, sections of these embryos are shown on the right side of each panel, in the inset, on the bottom right corner of every panel, the dotted line indicates the plane of sectioning. Embryos shown are at stage 11, vegetal views with dorsal up (**A**, **C**, **E**, and **G**) and at stage 13, dorsal views with anterior up (**B**, **D**, **F**, and **H**). *XMeis3* expression overlaps with dorsolateral expression of *Hoxd1*, *Hoxb4*, and *Hoxc6* in mesoderm at stage 11 (**A**, **C**, **E**, and **G**). *XMeis3* expression in ectoderm at stage 13 overlaps with expression of *Hoxd1* but not with expression of *Hoxb4* and *Hoxc6* (**B**, **D**, **F**, and **H**).

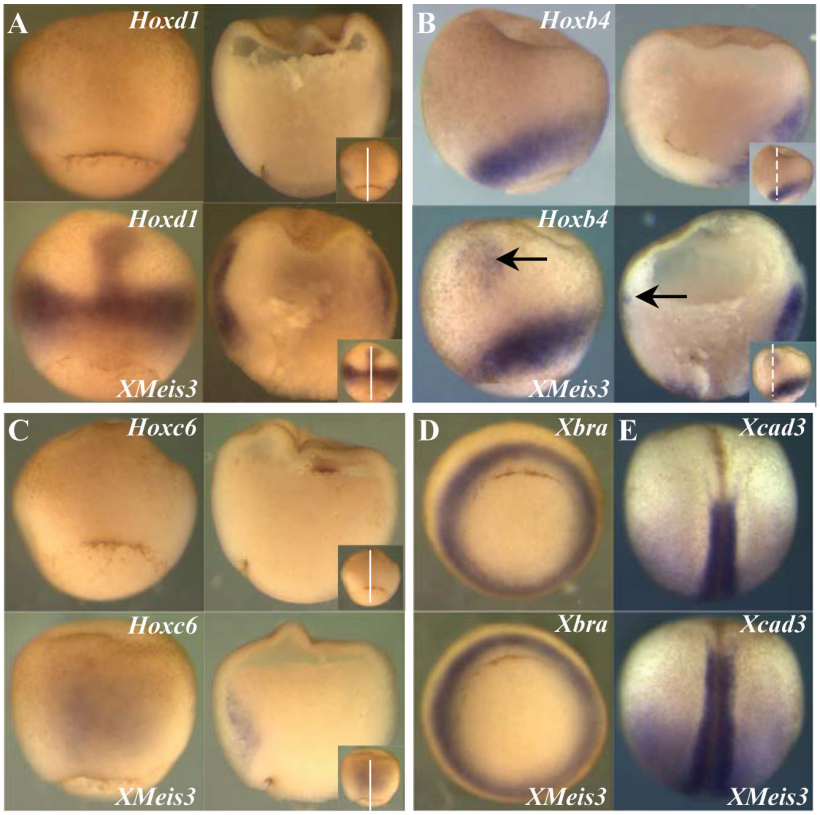

Figure 2. XMeis3 gain-of-function. Embryos were injected into the animal hemisphere of embryos at the one-cell stage with 2 ng synthetic mRNA containing the full-length coding region of *XMeis3*, and analysed by whole-mount *in situ* hybridisation. In each panel control embryos are shown on top, *XMeis3* injected embryos are shown on the bottom. For **A**, **B**, and **C**, whole mounts are shown on the left side, sections of these embryos are shown on the right hand of each panel. The plane of sectioning is depicted by the dotted line in the insets. **(A)** Expression of *Hoxd1*, whole mounts are shown in dorsal view, with anterior to the top, at stage 10.5. Lateral expression of *Hoxd1* in injected embryos is stronger and in a broader domain, the gap in expression on the dorsal mesoderm is closed and a streak of expression in dorsal mesoderm is observed. **(B)** Expression of *Hoxb4*, whole mounts are shown in lateral view, with dorsal to the left, at stage 11. Lateral expression of *Hoxb4* is not affected by injection of *XMeis3*, the black arrow points to a patch of ectopic expression in ectoderm. **(C)** Expression of *Hoxc6*, whole mounts are shown in dorsal view, with anterior to the top, at stage 10.5. Injected embryos show ectopic expression of *Hoxc6* in dorsal mesoderm, prior to initiation of endogenous expression of *Hoxc6*. **(D)** Expression of *Xbra*, embryos at stage 10.5 are shown in vegetal view with dorsal to the top. No change can be observed in the expression of the mesodermal marker *Xbra* as a result of injection of *XMeis3*. **(E)** Expression of *Xcad3*, embryos at stage 17 are shown in dorsal view with anterior to the top. The anterior expression boundary of the posterior marker *Xcad3* is shifted to a more anterior position by injection of *XMeis3*.

in *Hoxd1* expression, thus allowing possible synergistic effects to be observed. A downregulation of *Hoxd1* expression in embryos injected with a single morpholino and an additional reduction by injection of both morpholinos is visible (Fig. 6B). This suggests that indeed there is a synergistic effect of Hoxd1 and XMeis3 on establishment of *Hoxd1* expression in marginal zone mesoderm during gastrulation.

Discussion

XMeis3 expression overlaps early *Hox* expression

Much effort has been put into finding out details about the relation between *Hox* proteins and their cofactors Pbx/Exd and Meis/Hth. Although much has been accomplished, most of this work consists of *in vitro* binding studies. In *Xenopus* embryos, it has been shown that *XMeis3* has a function in hindbrain patterning (Salzberg *et al.*, 1999; Dibner *et al.*, 2001), these results are corroborated by recent reports concerned with Meis function in hindbrain formation in zebrafish embryos (Vlachakis *et al.*, 2001; Waskiewicz *et al.*, 2001; Choe *et al.*, 2002). We show here that *XMeis3* is expressed in marginal zone mesoderm significantly earlier than previously described (Salzberg *et al.*, 1999). We went on to show that an overlap in expression of *XMeis3* and early *Hox* genes is found in ventral and lateral mesoderm during gastrulation. In later phases of gastrulation the overlap is restricted to dorsolateral mesoderm. This co-localisation with early *Hox* genes suggests a role for *XMeis3* in the regulation of *Hox* gene expression in mesoderm during the early phases of gastrulation.

Ectopic *XMeis3* enhances *Hox* expression in mesoderm

By gain-of-function experiments we show that ectopic *XMeis3* is capable of inducing expression of *Hoxd1*, *Hoxb4*, and *Hoxc6*, expanding endogenous expression domains of these genes in mesoderm, and ectopically initiating expression in dorsal mesoderm. Interestingly, this induction of *Hox* expression by ectopic *XMeis3* can only be found as expansions of endogenous expression domains or in streaks of expression still in contact with the expanded endogenous domains of expression. This is most obvious for ectopic expression of *Hoxd1* in dorsal mesoderm, expanding into more animally located mesoderm and ectoderm. This suggests that ectopic *XMeis3* only enhances the expression of the assayed *Hox* genes, requiring factors already present in their endogenous *Hox* expression domains rather than

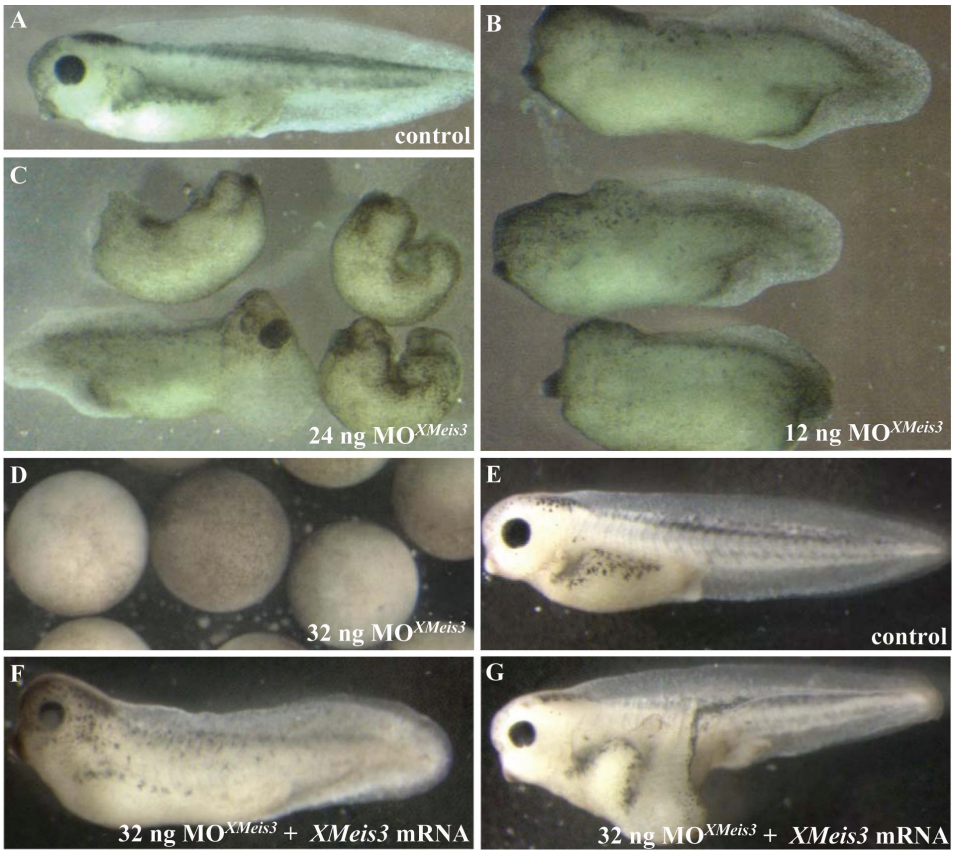

Figure 3. Effects of XMeis3 loss-of-function on the phenotype and the rescue of MO^{XMeis3}. Embryos at the one-cell stage were injected into the animal hemisphere with MO^{XMeis3} in amounts of 12 ng **(B)**, 24 ng **(C)**, and 36 ng **(D)**, and allowed to develop until the control embryos **(A)** reached tadpole stages. The specificity of MO^{XMeis3} is shown by the rescue with *XMeis3* synthetic mRNA. Embryos were injected with 32 ng of MO^{XMeis3} and 125 pg synthetic mRNA for *XMeis3* and allowed to develop until the control embryos reached the tad pole stage **(E)**, In the majority of the embryos a large part of the axis was rescued **(F)**, in a small number of embryos the phenotype could even be reversed, not only is the axis fully rescued but the embryo shown in **(G)** even possesses additional trunk structures as shown by the presence of somites in the axis outgrowth.

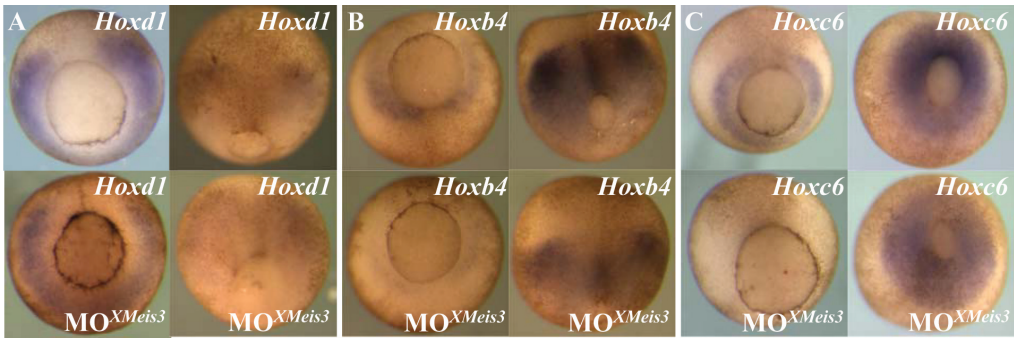

Figure 4. XMeis3 loss-of-function. Embryos were injected at the one-cell stage with 16 ng of the MO^{XMeis3}, and analysed by whole mount *in situ* hybridisation at stage 10.5/11, shown on the left side of each panel, and at stage 12, shown at the right side of each panel. Injected embryos are shown at the bottom of each panel, untreated embryos are shown on top. Shown are vegetal views with dorsal to the top. Expression of *Hoxd1* **(A)**, *Hoxb4* **(B)**, and *Hoxc6* **(C)** is downregulated in mesoderm of injected embryos at early gastrula stages. A reduction in ectodermal expression of the three studied *Hox* genes is observed in injected embryos at stage 12.

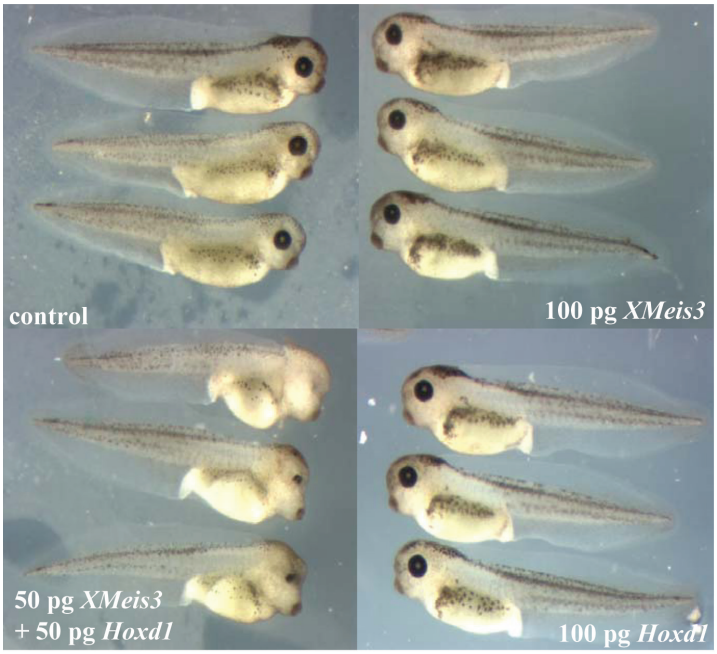

Figure 5. Synergistic effect between *Hoxd1* and *XMeis3* in ectopic expression. Embryos at the one-cell stage were injected into the animal hemisphere with either 100 pg *Hoxd1* mRNA, 100 pg *XMeis3* mRNA, or 50 pg of both mRNA's. A single injection of 100 pg of either factor is not sufficient to induce a phenotypic effect. The combination of half the amount of *Hoxd1* and *XMeis3*, results in posteriorisation, shown by a clear reduction of eye formation, and an anterior shift of the eye.

inducing expression *de novo*. These patterns are consistent with our idea that *XMeis3* enhances Hox autoregulation in mesoderm of *Xenopus* embryos.

XMeis3 is necessary for *Hox* expression in mesoderm and ectoderm

The injection of MO^{XMeis3} led to a downregulation of expression of all three *Hox* genes assayed. For *Hoxd1* and *Hoxb4* this held true for mesoderm and ectoderm, in the case of *Hoxc6*, mesodermal expression partially recovers during later phases of gastrulation, but ectodermal expression could not be observed. This indicates that *XMeis3* protein is necessary, in ventral and lateral mesoderm and in neurectoderm during gastrulation, for proper initiation and maintenance of *Hox* expression.
XMeis3 loss-of-function using small amounts of MO^{XMeis3} already led to a strong phenotype, indicating the necessity of XMeis3 function in anteroposterior patterning. This phenotype corroborates the results of Dibner and co-workers (2001). The sudden arrest in gastrulation at stage 11, caused by injecting a high amount of MO^{XMeis3} is very striking. We show by coinjecting a limited amount of *XMeis3* mRNA that the observed effect is not aspecific. The phenotype observed after injection of less morpholino, namely loss of trunk structures, head defects, and retarded tail formation described in this report and by Dibner and co-workers (2001), is therefore most likely a result of reduced XMeis3 function, not a complete loss of it. We cannot be certain that the phenotype caused by injection of 32 ng MO^{XMeis3} represents the complete loss-of-function phenotype, but it suggests the need for XMeis3 in two processes during early development: the progression of gastrulation and the subsequent patterning of the hindbrain.

Autoregulation by Hoxd1 is necessary for its establishment of expression in marginal zone mesoderm

Autoregulation dependent on Pbx/Exd has been shown for Hox paralog group 1 and 4 members (Pöpperl *et al.*, 1995; Gould *et al.*, 1997; Ryoo *et al.*, 1999; Ferretti *et al.*, 2000; Marty *et al.*, 2001; Streit *et al.*, 2002), this suggests that the shown regulation of *Hox* expression by XMeis3 could take place at the level of Hox autoregulation. Indeed, injection of MO^{Hoxd1} led to a reduction in *Hoxd1* expression. This is most likely the result of a reduction in *Hoxd1* translation, leading to a reduced amount of Hoxd1 protein. This reduction of Hoxd1 protein levels subsequently led to an apparent reduction in *Hoxd1* transcription. This suggests that Hoxd1 autoregulation is an essential step in the establishment, and not only the maintenance, of *Hoxd1* expression in

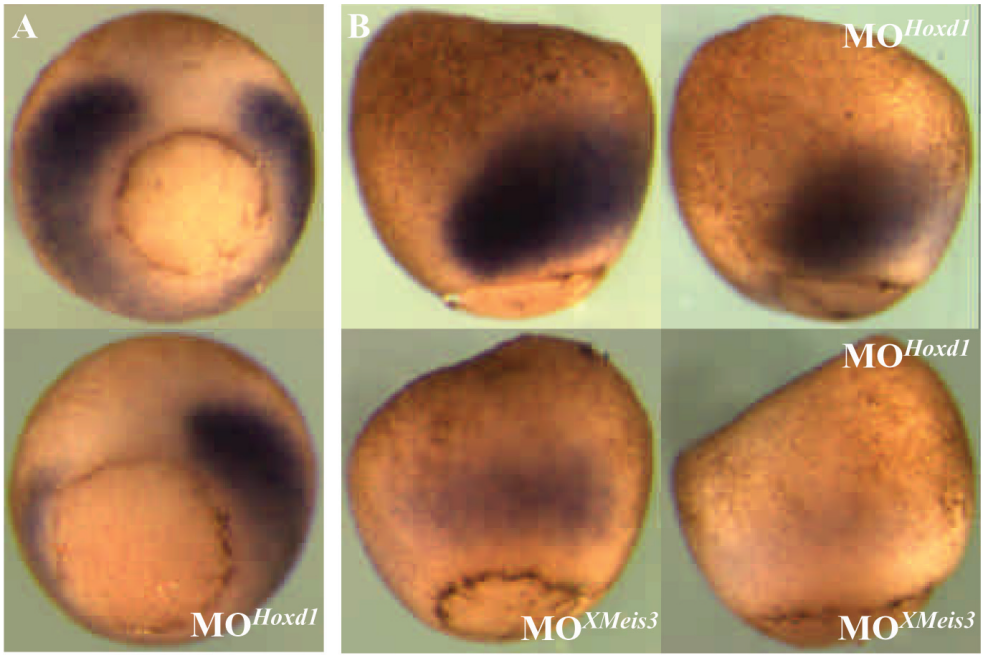

Figure 6. Synergistic effects in loss-of-function of Hoxd1 and XMeis3. **(A)** Embryos were injected with 32 ng of MO^{Hoxd1} into the lateral marginal zone on the left side of embryos, rendering the un-injected side an internal control. Embryos were allowed to develop until control stage 11 and assayed by *in situ* hybridisation for expression of *Hoxd1*. Embryos are shown in vegetal view, with dorsal up. Expression of *Hoxd1* is reduced on the left side of injected embryos (shown on the bottom of the panel). **(B)** To investigate whether there is synergy between Hoxd1 and XMeis3, 16 ng MO^{XMeis3} and 16 ng MO^{Hoxd1} were injected, together and separately, into the animal hemisphere of one-cell stage embryos, at stage 11 the embryos were harvested and assayed for expression of *Hoxd1* by *in situ* hybridisation. Embryos are shown in lateral view, with dorsal to the left. Injection of MO^{Hoxd1} and MO^{XMeis3} separately leads to a reduction in expression of *Hoxd1*, the co-injection leads to a reduction as compared to injection of either MO^{XMeis3} or MO^{Hoxd1} separately. This suggests that Hoxd1 and XMeis3 work synergistically in establishment of *Hoxd1* expression in mesoderm during early gastrula stages.

mesoderm during gastrulation in *Xenopus* embryos. The observed reduction of *Hoxd1* expression could also be explained if binding of MO^{Hoxd1} to mRNA led to a reduction in stability of the messenger, however this potential effect has, to our knowledge, never been reported. The necessity for *Hoxd1* autoregulation in mesoderm is a remarkable discovery considering that vertebrate Meis family members have so far only been shown to be involved in Hox autoregulation in the hindbrain. Here we show that XMeis3 is an essential factor for establishment of stable *Hoxd1* expression in marginal zone mesoderm. A second noteworthy aspect is that apparently Hoxd1 loss-of-function is not fully, if at all, rescued by the other labial type gene normally expressed during the early phases of gastrulation, *Hoxa1* and *Hoxb1* as would be expected from a viewpoint of redundancy. Either Hoxa1 and Hoxb1 are not capable of inducing the expression of *Hoxd1*, which seems unlikely taking into account the redundant function of these paralog group members (reviewed in Morrison, 1998, and references therein), or expression of *Hoxa1* and *Hoxb1* is also reduced or prevented by *Hoxd1* loss-of-function, this would suggest the necessity of Hoxd1 to induce the two other labial homologous during gastrulation in *Xenopus* embryos. Additional experiments are needed to distinguish between the two possibilities but whatever the outcome, this sheds new light on the initiation and establishment of expression of the first *Hox* genes of the *Hox* cascade.

Synergy between *Hoxd1* and *XMeis3*

The synergistic effects we have observed in the gain-of-function experiment by injection of synthetic *XMeis3* and *Hoxd1* mRNA together show that indeed these two factors, when co-expressed can generate a phenotype that cannot be accomplished by injecting double the amount of either factor separately. These results corroborate the findings of Vlachakis and co-workers (2001), who have shown that in zebrafish embryos, Meis3, Pbx4, and Hoxb1 synergise to promote hindbrain fate. The combined Hoxd1 and XMeis3 loss-of-function supports the suggested synergy, while sub optimal amounts of either morpholino against *Hoxd1* or *XMeis3* led to a reduction of the *Hoxd1* expression, the combination led to a significantly stronger reduction. This adds to the evidence for a synergistic relation between Hoxd1 and XMeis3. Taken together our results show that XMeis3 is necessary in marginal zone mesoderm to establish the expression of early Hox genes. This XMeis3-mediated mesodermal *Hox* cascade is of vital importance for axis formation and AP patterning.

Materials and methods

Xenopus embryos and microinjections

Pigmented *Xenopus laevis* embryos were obtained by *in vitro* fertilisation, and after dejelling in a 2% cysteine solution (pH 8.0), cultured in 0.1x Marc's Modified Ringers's (MMR) (Sive *et al.*, 2000), containing 50 µg/ml gentamycin at 14-21 ^{o}C. Embryos were injected in 1x MMR + 4% ficoll and afterwards transferred to 1x MMR + 1% Ficoll, and cultured in this medium for 1 to 7 hours, after which they were transferred and to 0.1x MMR in which they were cultured until harvesting. Staging of the embryos was performed according to Nieuwkoop and Faber (Nieuwkoop and Faber, 1967). Embryos at the one-cell stage were injected into the animal pole with synthetic mRNA dissolved in water. The synthetic capped mRNA was made using the Ambion mMessage mMachine Kit with CS2-*XMeis3*, or CS2-*Hoxd1*, linearised with *Not*I, as template. CS2-*XMeis3* was constructed by cloning the full-length coding region of *XMeis3*, obtained by PCR using stage 15 cDNA as template and the following primers: f: 5'-gcgggatccatggcacaaaggtatgatgag, r: 5'–cgcctcgagcatgtagtgccactgcccctcc, containing an *BamH*I or a *Xho*I restriction site respectively, in the CS2+ vector (Rupp *et al.*, 1994) using the restriction sites in the primers. CS2-Hoxd-1 contains the complete coding sequence of *XHoxd1* in CS2+, kindly provided by W. Van den Akker.
MO^{XMeis3}, supplied by Gene Tools, LLC, has the sequence: 5'-cctttgtgccattccgagttgggtc, and was injected in amounts of 8 to 48 ng in a concentration of 8 ng/nl. MO^{contr}, supplied by Gene Tools, LLC, has the sequence: 5'-cctcttacctcagttacaatttata and was injected using the same amounts and concentrations as MO^{XMeis3}.

Whole mount *in situ* hybridisation and antisense probes

Whole mount *in situ* hybridisations were performed according to Harland (1991), with minor modifications. The antisense RNA probes were generated by run off *in vitro* translation using DIG RNA labelling mix (Roche), and T7 or Sp6 RNA polymerase (Promega). The probes were generated using the following templates: *Hoxd1*: (Sive and Cheng, 1991), *Hoxb4*: a 708 bp fragment containing the complete *Hoxb-4* ORF cloned in pGEMTE, *Hoxc6*: a 998 bp *Hoxc-6* fragment in pGEM1 containing a part of the homeodomain and extending into the 3' UTR, *Xcad3*: (Pownall *et al.*, 1996); *Xbra*: pSP73Xbra (Smith *et al.*, 1991).

Acknowledgments

We thank W. van den Akker for kindly providing the CS2-*Hoxd1* plasmid prior to publication and C. McNulty, G. Mainguy, and S. Wacker for critically reading the manuscript. PIdR was supported by a NWO-grant and EU "HPRN" and "Quality of life" programs.

References

Abu-Shaar, M., Ryoo, H. D., and Mann, R. S., (1999), Control of the nuclear localization of Extradenticle by competing nuclear import and export signals. Genes Dev. 13: 935-945.

Bürglin T. R. and Ruvkun G., (1993), The *Caenorhabditis elegans* homeobox gene cluster. Curr. Opin. Genet. Dev. 3: 615-620.

Bürglin T. R., (1997), Analysis of TALE superclass homeobox genes (Meis, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals. Nucleic Acid Res. 25: 4173-4180.

Bürglin, T. R., Ruvkun, G., Coulson, A., Hawkins, N. C., McGhee J. D., Schaller, D., Wittmann, C., Müller, F., Waterston, R. H., (1991), Nematode homeobox cluster. Nature 351: 703.

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 & Dev. 9: 663-674.

Choe, S. K., Vlachakis, N., and Sagerström, C. G., (2001), Meis family proteins are required for hindbrain development in the zebrafish. Development 128: 585-595.

De Robertis, E. M., Oliver, G., and Wright, C. V. E., (1989), Determination of Axial Polarity in the vertebrate Embryo: Homeodomain Proteins and Homeogenetic Induction. Cell 57: 189-191.

Deschamps, J., van den Akker, E., Forlani, S., de Graaff, W., Oosterveen, T., Roelen, B., and Roelfsema, J., (1999), Initiation, establishment and

maintenance of *Hox* gene expression patterns in the mouse. Int. J. Dev. Biol. 43: 635-650.

van Dijk, A., and Murre, C., (1994) extradenticle Raises the DNA Binding Specificity of Homeotic Selector Gene Products. Cell 78:617-624.

Dibner, C., Elias, S., and Frank, D, (2001), XMeis3 protein is required for proper hindbrain patterning in *Xenopus laevis* embryos. Dev 128: 3415-3426.

Duboule, D., and Dolle, P., (1989), The structural and functional organization of the murine HOX gene family resembles that of Drosophila homeotic genes. EMBO J. 8: 1497-1505.

Ferretti, E., Marshall, H., Pöpperl, H., Maconochie, M., Krumlauf, R., and Blasi, F., (2000), Segmental expression of Hoxb2 in r4 requires two separate sites that integrate cooperative interactions between Prep1, Pbx and Hox proteins. Development 127, 155-166.

Gaunt, S. J., and Strachan, L. (1996), Temporal colinearity in expression of anterior Hox genes in developing chick embryos. Dev. Dyn. 207: 270-280.

Gould, A., Morrison, A., Sproat, G., White, R. A., and Krumlauf, R., (1997), Positive cross-regulation and enhancer sharing: two mechanisms for specifying overlapping *Hox* expression patterns. Genes Dev. 11: 900-913.

Graham, A., Papalopulu, N., and Krumlauf, R., (1989), The murine and Drosophila homeobox gene complexes have common features of organization and expression. Cell 57: 367-378.

Grieder, N. C., Marty, T., Ryoo, H. D., Mann, R. S., Affolter, M., (1997), Synergistic activation of a Drosophila enhancer by HOM/EXD and DPP signaling. EMBO J. 16: 7402-7410.

Heasman, J., (2002), Morpholino Oligos: Making Sense of Antisense? Dev. Biol. 243: 209-214.

Harvey, R. P., Melton, D. A., (1988), Microinjection of synthetic Xhox-1A homeobox mRNA disrupts somite formation in developing Xenopus embryos. Cell 1988 53: 687-697.

Heasman, J., (2002), Morpholino Oligos: Making Sense of Antisense? Dev. Biol. 243: 209-214.

Jabet, C., Gitti, R., Summers, M. F., and Wolberger, C., (1999), NMR Studies of the Pbx1 TALE Homeodomain Protein Free in Solution and Bound to DNA: Proposal for a Mechanism of HoxB1-Pbx1-DNA Complex Assembly. J. Mol. Biol. 291: 521-530.

Jaw, T. J., You, L. R., Knoepfler, P. S., Yao, L. C., Pai, C. Y., Tang, C. Y., Chang, L. P., Berthelsen, J., Blasi, F., Kamps, M. P., Sun, Y. H., (2000), Direct interaction of two homeoproteins, Homothorax and Extradenticle, is essential for EXD nuclear localization and function. Mech. Dev. 91: 279-291.

Knoepfler, P. S., and Kamps, M. P., (1995), The Pentapeptide Motif of Hox Proteins Is Required for Cooperative DNA Binding with Pbx1, Physically Contacts Pbx1, and Enhances DNA binding by Pbx1. Moll. Cell. Biol. 15: 5811-5819.

Knoepfler, P. S., Bergstrom, D. A., Uetsuki, T., Dac-Korytko, I., Sun, Y. H., Wright, W. E., Tapscott, S. J., and Kamps, M. P., (1999), A conserved motif N-terminal to the DNA-binding domains of myogenic bHLH transcription factors mediates cooperative DNA binding with Pbx-Meis1/Prep1. Nucleic Acids Res., 27: 3752-3761.

Kolm, P. J. and Sive, H. L., (1995), Regulation of the Xenopus labial Homeodomain Genes, HoxA1 and HoxD1: Activation by Retinoids and Peptide Growth Factors. Dev. Biol. 167: 34-49.

Lawrence, P. A. and Morata, G., (1994), Homeobox genes: their function in Drosophila segmentation and pattern formation. Cell 78: 181-189.

Manak, J. K. and Scott, M. P. (1994), A class act: conservation of homeodomain protein functions. Development Suppl., 61-77.

Marty, T., Vigano, M. A., Ribeiro, C., Nussbaumer, U., Grieder, N. C., Affolter, M. A., (2001), HOX complex, a repressor element and a 50 bp sequence confer regional specificity to a DPP-responsive enhancer. Development 128: 2833-2845.

McGinnis, W. and Krumlauf, R., (1992), Homeobox genes and axial patterning. Cell 68: 283-302.

Morrison, A. D., (1998), 1 + 1 = r4 and much much more. Bioessays 20: 794-797.

Nieuwkoop, P. D. and Faber, J., (1967), Normal table of *Xenopus laevis* (Daudin). Amsterdam: North Holland.

Neuteboom, S. T. C. and Murre, C., (1997), Pbx raises the DNA binding specificity but Not the Selectivity of Antennapedia Hox Proteins. Moll. Cell. Biol.17: 4696-4706.

Oliver, G., Wright, C. V., Hardwicke, J., De Robertis, E. M., (1988), Differential antero-posterior expression of two proteins encoded by a homeobox gene in Xenopus and mouse embryos. EMBO J. 7: 3199-3209.

Passner, J. M., Ryoo, H. D., Shen, L., Mann, R. S., and Aggarwal, A. K., (1999) Structure of a DNA-bound Ultrabithorax-Extradenticle homeodomain complex. Nature 397: 714-719.

Peltenburg, L. T. C., and Murre, C., (1996), Engrailed and Hox homeodomain proteins contain a related Pbx interaction motif that recognizes a common structure present in Pbx. EMBO J. 15: 3385-3393.

Pinsonneault, J., Florence, B., Vaessin, H., and McGinnis, W., (1997) A model for extradenticle function as a switch that changes HOX proteins from repressors to activators. EMBO J. 16:2032-2042.

Piper, D. E., Batchelor, A. H., Chang, C., Cleary, M. L., Wolberger, C., (1999), Structure of a HoxB1-Pbx1 Heterodimer Bound to DNA: Role of the Hexapeptide and a Fourth Homeodomain Helix in Complex Formation. Cell 96: 587-597.

Pöpperl, H., Bienz, M., Struder, M., Chan, S. K., Aparicio, S., Brenner, S., Mann, R. S., and Krumlauf, R., (1995), Segmental expression of *Hoxb1* is Controlled by a Highly conserved Autoregulatory Loop Dependent upon *exd*/*pbx*. Cell 81: 1031-1042.

Pownall, M. E., Tucker, A. S., Slack, J. M. W., and Isaacs, H. V., (1996), eFGF, Xcad3 and Hox genes form a molecular pathway that establishes the anteroposterior axis in *Xenopus*. Development 122: 3881-3892.

Rupp, R. A. W., Snider, L., Weitraub, H., (1994), *Xenopus* embryos regulate the nuclear localisation of XMyoD. Genes Dev. 8: 1311-1323.

Ryoo, H. D., and Mann, R. S., (1999), The control of trunk Hox specificity and activity by Extradenticle. Genes & Dev. 13: 1704-1716.

Ryoo, H. D., Marty, T., Casares, F., Affolter, M., Mann, R. S., (1999) Regulation of Hox target genes by a DNA bound Homothorax/Hox/Extradenticle complex. Development 126: 5137-5148.

Salzberg, A., Elias, S., Nachaliel, N., Bonstein, L., Henig, C., Frank, D., (1999), A Meis family member protein caudalizes neural cell fates in *Xenopus*. Mech. Dev. 80: 3-13.

Sive, H. L., and Cheng, P. F., (1991), Retinoic acid perturbs the expression of *Xhox.lab* genes and alters mesodermal determination in *Xenopus laevis*. Genes Dev. 5: 1321-1332.

Sive, H. L., Grainger, R. M., Harland, R. M., (2000), Early development of *Xenopus laevis* A laboratory manual. Cold Spring Harbor Laboratory press.

Smith, J. C., Price, B. M., Green, J. B., Weigel, D., and Herrmann, B. G., (1991), Expression of a *Xenopus* homolog of *Brachyury* (T) is an immediate-early response to mesoderm induction. Cell. 67: 79-87.

Streit, A., Kohler, R., Marty, T, Belfiore, M., Takacs-Vellai, K., Vigano, M. A., Schnabel, R., Affolter, M., and Müller, F., (2002), Conserved Regulation of the Caenorhabditis elegans labial/Hox1 gene ceh-13. Dev. Biol. 232: 96-108.

Studer, M., Gavalas, A., Marshall, H., Ariza-McNaughton, L., Rijli, F. M., Chambon, P., Krumlauf, R., (1998), Genetic interactions between Hoxa1 and Hoxb1 reveal new roles in regulation of early hindbrain patterning. Development 125: 1025-1036.

Vlachakis, N., Choe, S. K., and Sagerström, C. G., (2001), Meis3 synergizes with Pbx4 and Hoxb1b in promoting hindbrain fates in the zebrafish. Development 128: 1299-1312.

Wacker, S. A., Jansen, H. J., McNulty, C. L., Houtzager, E., and Durston, A. J., Timed interactions between a Hox cascade and the Spemann organiser generate positional information during vertebrate gastrulation. Submitted.

Waskiewicz, A. J., Rikhof, H. A., Hernandez, R. E., and Moens, C. B., (2001), Zebrafish Meis functions to stabilize Pbx proteins and regulate hindbrain patterning. Development 128: 4139-4151.

Chapter 6

Hexapeptide revisited

Paul M. J. In der Rieden, Gaëll Mainguy,
Joost Woltering, and Antony J. Durston

Hubrecht Laboratorium,
Nederlands Instituut voor Ontwikkelingsbiologie,
Uppsalalaan 8, 3584 CT, Utrecht, The Netherlands.

Introduction

The ParaHox cluster, an evolutionary sister of the Hox clusters, contains three genes, in chordates called *Gsh*, *Pdx1*, and *Cdx* (or *Cad*) respectively (Brooke *et al.*, 1998; Coulier *et al.*, 2000). An intact ParaHox cluster, containing the genes *Gsh1*, *Pdx1* and *Cdx2* is found in human and mouse, on chromosome 5 or 13, respectively. The other ParaHox genes, *Gsh2*, *Cdx1*, and *Cdx4* are not, or no longer, organised in a cluster, as they are all located on different chromosomes (Pollard and Holland, 2000), although it cannot be excluded that the genes were duplicated independently of complete cluster duplications. Amino acid sequence comparison studies using the homeodomains of the Hox and ParaHox proteins has revealed that Gsh, Pdx1, and Cdx/Cad proteins are most similar to Hox paralog groups 1 and 2, 3, and 9 through 13, respectively (Kourakis and Martindale, 2000; Yanze *et al.*, 2001). However the homeodomains of Cdx1 and Cdx2 are closer to Hox paralog groups 8 and 9, and even to groups 1 and 2 than to the most posterior group (Van den Akker *et al.*, 2002).
Studies concerning Hox target specificity have revealed that PBC-class co-factors, subfamily of the TALE-class of homeodomain proteins, are employed by most Hox proteins to enhance binding specificity and affinity (Pöpperl *et al.*, 1995; Mann and Chan, 1996; Di Rocco *et al.*, 1997). The Hox/PBC interaction is mediated via a so-called hexapeptide motif. NMR and X-ray crystallographic analysis of Lab/Exd and Hoxb1/Pbx1 fragments bound to a short DNA sequence has revealed that the tryptophan residue in the core of the hexapeptide binds a pocket formed by the atypical homeodomain of PBC family members (Jabet *et al.*, 1999; Passner *et al.*, 1999; Piper *et al.*, 1999). This pocket is composed of the three amino acid loop extension of the PBC homeodomain, residues in the third helix of the homeodomain, and a residue in the C-terminal helix of PBC homeodomains (Piper *et al.*, 1999). From known sequences the hexapeptide has been defined as a tryptophan residue surrounded by hydrophobic residues that is flanked by a lysine or arginine residue two to five amino acids C-terminally (Knoepfler *et al.*, 1999). Previously we have shown that highly conserved amino acids are clustered around the hexapeptide sequence, and consistently identify Hox proteins as belonging to a particular paralog group (Morgan *et al.*, 2000). We and others have suggested a recognition code. In contrast, the mechanism(s) in which ParaHox proteins achieve target specificity are largely unknown. Recently, the ParaHox gene *Pdx1* has been shown to depend on Pbx1 to fully employ its function in pancreatic development (Kim *et al.*, 2002), raising the possibility that the "code" could be extended to

ParaHox genes. In addition, it has been reported that Cdx proteins contain a hexapeptide sequence as well (Van den Akker *et al.*, 2002). To gain more insight into ParaHox/TALE-class co-factor interaction we undertook a search for hexapeptide sequences in all known ParaHox proteins, and if so whether conservation of flanking sequence can also be found in the ParaHox proteins. We found that all described Cdx and Pdx1 members contain a hexapeptide sequence and that in these factors hexapeptide-flanking sequences conservation exists. In contrast, Gsh members do not possess a hexapeptide sequence. More generally we searched for the presence of a hexapeptide sequence in all of the members of the Antp-class of homeodomain proteins, and found them to be widely distributed. This suggests a functional interaction between Antp-class homeodomain proteins and TALE-class co-factors early during evolution.

Sequences used in this study

Gsh: *Ciona intestinalis* Gsx (AF305500), Human Gsh1 (AB044158), Human Gsh2 (AB028838), Mouse Gsh1 (NM_008178), Mouse Gsh2 (S79041), *Nephasoma minuta* Gsx (AF363231), *Oryzias latipes* Gsh1 (AF035573), *Phascolion strombi* Gsx (AF363230), and *Podocoryne carnea* Gsx (AF268446). **Pdx1**: Human Pdx-1 (U35632), *Mesocricetus auratus* Pdx1 (U73854), Mouse Pdx-1 (XM_124700), Rat STF-1 (S67435), *Xenopus laevis* XlHbox8 (X16849), and Zebrafish Pdx1 (NM_131443). **Cdx**: *Anopheles gambiae* Cdx (AF119382), *Bombyx mori* Cdx (D16683), Carp Cdx1 (X80668), *Caenorhabditis elegans* Cad (NM_065590), *Drosophila melanogaster* Cad (NM_134301), Zebrafish Cad1 (NM_131109), Chicken CdxA (AB046532), Chicken CdxB (AF353624), Chicken CdxC (U80614), *Halocynthia roretzi* Cad (AB031032), *Herdmania curvata* Cdx (AF242305), Human Cdx1 (U51095), Human Cdx2 (NM_001265), Human Cdx4 (NM_005193), *Mesocricetus auratus* Cdx3 (X81404), Mouse Cdx1 (BC019986), Mouse Cdx2 (NM_007673), Mouse Cdx4 (L08061), Rat Cdx2 (NM_023963), *Tribolium castaneum*, (AJ005421), *Xenopus laevis* Cad2 (U04302), *Xenopus laevis* Cad3 (U02034), *Xenopus tropicalis* Cad1 (AF417197), *Xenopus tropicalis* Cad2 (AF417198), and *Xenopus tropicalis* Cad3 (AF417199). **HB9**: Chicken HB9 (AF066861), *Fugu rubripes* (SINFR UG00000071220) Human HB9 (U07664), Mouse HB9 (NM_019944) and *Xenopus laevis* XHB9 (AF072382). **Msx**: *Ambystoma mexicanum* Msx1 (BAA11574), *Ambystoma mexicanum* Msx2 (AAD28493), *Bos taurus* Msx1 (BAA20367), *Canis familiaris* Msx2 (AJ277753), Chicken Ghox7 (D10372), Chicken Msx2 (S64478), Human Msx1 (M97676), Human Msx2 (S75361),

Mouse Msx1 (BC016426), Mouse Msx2 (L11739), and Rat Msx1 (D83036). **Nkx**: *Fugu rubripes* Nkx6 (SINFRUG00000085230), *Fugu rubripes* Nkx6-1 (SINFRUG00000050033), Human Gtx (AF184215), Human Nkx6-1 (NM_0 06168), *Mesocricetus auratus* Nkx6-1 (CAA57166), Mouse Gtx (ENSMUSG 00000041309), Mouse Nkx6 (ENSMUSG00000035187), and Rat Nkx6-1 (AF004431), **Emx**: *Fugu rubripes* Emx1 (SINFRUG00000074711), *Fugu rubripes* Emx2 (SINFRUG00000060256), Human Emx1 (ENSG0000013563 8), Human Emx2 (AF301598), *Oryzias latipes* Emx1 (AJ250402), *Oryzias latipes* Emx2 (AJ132403), Zebrafish Emx1 (D32214), Zebrafish Emx2 (D32215), **Hmx**: Human Hmx1 (M99587) and Mouse Hmx1 (AF009367). **Tlx**: Chicken Tlx1 (AF071874), Chicken Tlx3 (AF071875), Human Tlx1 (M62626), Human Tlx2 (BC006356), Mouse Tlx1 (S70632), Mouse Tlx2 (M75953), Mouse Tlx3 (AJ223801), *Xenopus laevis* XHox11 (AF283694), *Xenopus laevis* Xhox11L2 (AF283693), and zebrafish Tlx3 (AY045753).

Results and discussion

The analysis of the complete protein sequences of Gsh members (all sequences used in this study and accession numbers are listed at the end) revealed that none of the Gsh proteins contains a hexapeptide, or a derived sequence. Therefore it is highly unlikely that they use PBC-class proteins as co-factors to enhance target specificity and/or affinity.

Human Pdx1 was shown to contain a hexapeptide (Goudet *et al.*, 1999). Sequence alignments of the Pdx1 homologs revealed a conservation of the hexapeptide sequence in vertebrates (Fig. 1A). Sequence comparison of the homeodomain and 13 amino acids C-terminal to the homeodomain showed that Pdx1 members resemble closely the third Hox paralog group (Kourakis and Martindale, 2000), accordingly the conserved hexapeptide of Pdx1 family members resembles mostly the hexapeptide of the third Hox paralog group (FPWMK, Morgan *et al.*, 2000). In addition, sequence conservation of the flanking regions can also be observed (Fig. 1A). The conservation of the sequences flanking the Pdx1 hexapeptide is in corroboration with the sequence conservation found in comparable sequences in Hox paralog group proteins (Morgan *et al.*, 2000).

Phylogenetic analysis of the complete amino acid sequences of Cad/Cdx homologs reveals a division of the vertebrate caudal members into 3 groups (Fig. 2), that we named Cdx1, Cdx2, and Cdx3. We suggest reclassification of the Cdx/Cad members accordingly. A hexapeptide sequence can be found in all Cdx proteins (Fig 1B).

Because sequence conservation of hexapeptides and their flanking regions per paralog group is a feature described for Hox paralog groups 1 to 8

A Pdx1 ggLEepnRvqL**PFPWMK**STKaHaWKgQWagGaY
B Cdx1 SPxAqRrx**pYEWMR**Rs---x_{7-10}--GKTRT
Cdx2 QLSPxGqRRx**lcEWmR**KPaq-x_{6-7} --vKTRT
Cdx3 SPxxxrxs**sYeWMK**TVqst-----GKTRT
C HB9 DPIKlsAgTFQ**LDqWLR**aSTAGMiLPKMpDF
D Msx1 AESPdkpe**RtPWMQ**sPrFSPPpaRRLSPP
Msx2 ASVKSEnsED**GaaWIQ**epGRYSPPPRHlSPt
Nkx PIFWPGVmQs**pPWRD**ARLA
Hmx SDRDSPETGEEMGRA**EgAWPR**GPG
Emx FxsqqRDplt**FYPWVl**hRyrylghRFQ
Tlx ltgl**tFPWme**SsRRfxKdRfT

Figure 1. Conserved hexapeptide sequences and flanking regions. Capital letters denote conserved amino acids, in lower case predominant amino acids are depicted, and 'x' indicates that that the amino acid at that position is not conserved (based on standard IUB codes). The amino acids that define the hexapeptide are shown in bold. **(A)** Pdx1 hexapeptide and flanking sequence conservation, the sequences compared are Human Pdx-1, *Mesocricetus auratus* Pdx1, Mouse Pdx-1, Rat STF-1, *Xenopus laevis* XlHbox8, and Zebrafish Pdx1. **(B)** Conserved hexapeptide and flanking region of the Cdx groups. The putative phosphorylation motif S-P is underlined. The sequences compared are for each Cdx group as follows. Cdx1: Chicken, Human, Mouse, *Xenopus laevis*, and *Xenopus tropicalis*. Cdx2: Chicken, Human, *Mesocricetus auratus*, Mouse, Rat, *Xenopus tropicalis*. Cdx3: Carp, Chicken, Human, Mouse, *Xenopus laevis*, and *Xenopus tropicalis*, Zebrafish. **(C)** Hexapeptide and flanking sequence conservation of Chicken HB9, *Fugu rubripes* HB9, Human HB9, Mouse HB9, and *Xenopus laevis* XHB9. **(D)** Putative hexapeptide and flanking sequence conservation in NKL genes. Msx: *Ambystoma mexicanum* Msx1 and Msx2, *Bos taurus* Msx1, *Canis familiaris* Msx2, Chicken Ghox7 and Msx2, Human Msx1 and Msx2, Mouse Msx1 and Msx2, and Rat Msx1. Nkx: *Fugu rubripes* Nkx6 and Nkx6-1, Human Gtx and Nkx6-1, *Mesocricetus auratus* Nkx6-1, Mouse Gtx and Nkx6, and Rat Nkx6-1. Emx: *Fugu rubripes* Emx1 and Emx2, Human Emx1 and Emx2, *Oryzias latipes* Emx1 and Emx2, and Zebrafish Emx1 and Emx2. Hmx: Human Hmx1 and Mouse Hmx1. Tlx: Chicken Tlx1 and Tlx3, Human Tlx1 and Tlx2, Mouse Tlx1, Tlx2, and Tlx3, *Xenopus laevis* Xhox11 and Xhox11L2, and Zebrafish Tlx3.

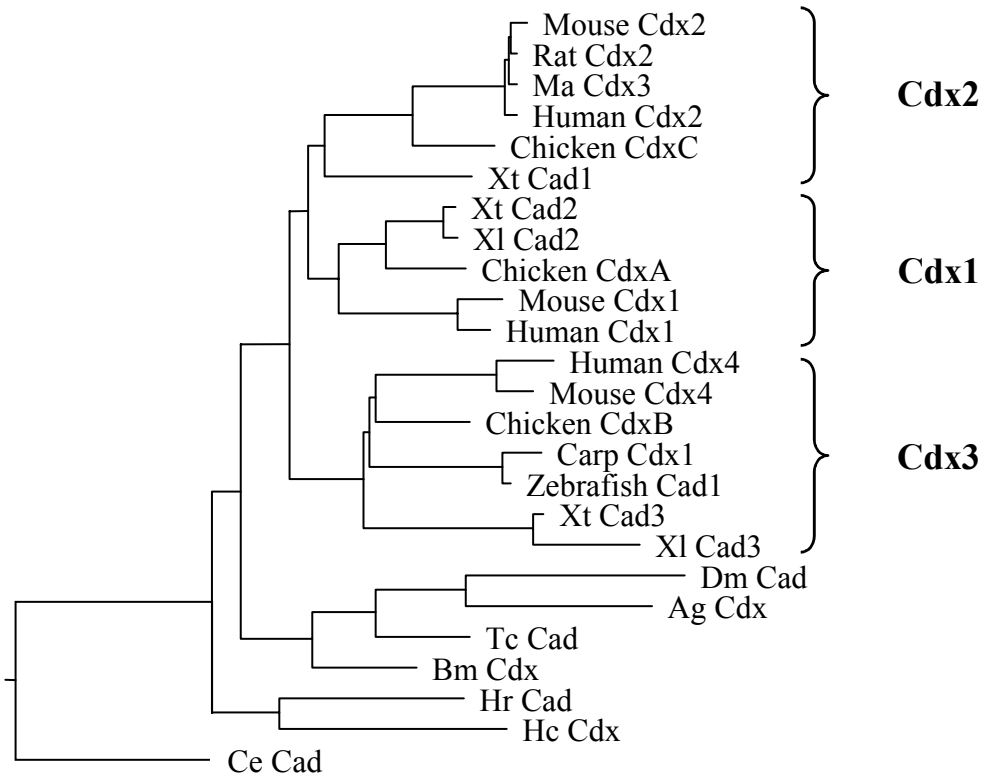

Figure 2. Phylogenetic tree of complete Cdx/Cad proteins, constructed using the sequences and abbreviations as follows: *Anopheles gambiae* (Ag) Cdx; *Bombyx mori* (Bm) Cdx; Carp Cdx1; *Caenorhabditis elegans* (Ce) Cad; *Drosophila melanogaster* (Dm) Cad; Zebrafish Cad1; Chick CdxA; Chick CdxB; Chick CdxC; *Halocynthia roretzi* (Hr) Cad; *Herdmania curvata* (Hc) Cdx; Human Cdx1; Human Cdx2; Human Cdx4; *Mesocricetus auratus* (Ma) Cdx3; Mouse Cdx1; Mouse Cdx2; Mouse Cdx4; Rat Cdx2; *Tribolium castaneum* (Tc); *Xenopus laevis* (Xl) Cad2; *Xenopus laevis* (Xl) Cad3; *Xenopus tropicalis* (Xt) Cad1; *Xenopus tropicalis* (Xt) Cad2; *Xenopus tropicalis* (Xt) Cad3.

(Morgan *et al.*, 2000), and now also for Pdx1 ParaHox members, we analysed the sequences of the hexapeptides and flanking region of the different Cdx groups. The hexapeptide sequences of the three different Cdx groups are highly conserved (Fig. 1B), and distinct from other described hexapeptides. When the flanking sequences of the Cdx hexapeptides are compared to each other within each group, sequence conservation defining an individual Cdx group can also be observed (Fig. 1B).

Despite the differences mentioned above between the Cdx groups, all members of the Cdx family share a number of characteristics. Firstly, in all members of the Cdx1 and Cdx2 groups an acidic amino acid residue is present directly N-terminal to the tryptophan in the core sequence of the hexapeptide (Fig. 1B). In the Cdx3 group the Cad3 proteins of *Xenopus laevis* and - *tropicalis* also contain an acidic residue at that position. This acidic residue most likely influences the target specificity of the Cdx proteins in dimers with Pbx co-factors, and the fact that it is also found in all the known insect Cdx homologs suggests that this is an ancient characteristic of the Cdx proteins. Comparison of the Cdx hexapeptide to the other known hexapeptides reveals that the labial group Hox proteins have an acidic residue at the same position as the Cdx proteins (Morgan *et al.*, 2000). Secondly, the Cdx/Cad hexapeptides contain an arginine or lysine residue at positions 2 and 3 (the latter with exception of Cdx3 proteins), C-terminal to the core tryptophan residue (Fig. 1B); this feature is only found in the hexapeptides of the fourth Hox paralog members (Morgan *et al.*, 2000). Thirdly, an arginine or lysine residue is found at the –5 position to the core tryptophan residue in all known insect and ascidian Cdx homologs and an arginine at position –6 (and often also –5) in Cdx1, Cdx2, and Cdx3 groups, with the exception of the *Xenopus laevis* and – *tropicalis* Cad3 proteins, which have a serine residue at the –6 position instead of a basic residue (Fig. 1B). This is a feature shared with the Hox paralog group 3 proteins, where in the flanking sequences of the hexapeptide a conserved lysine is found at a position –5 to the core tryptophan residue (Morgan *et al.*, 2000). Fourthly, in all described Cdx/Cad proteins, except those of *C. elegans* and cnidarians, an S-P sequence is found N-terminally to the hexapeptide (Fig. 1B). In the vertebrate Cdx proteins the serine residue is found at position -11 relative to the core tryptophan residue of the hexapeptide. In the insect Cad/Cdx proteins the S-P sequence is found between positions –4 to –9, depending on the species. We analysed the sequence flanking the serine residue (10 amino acids C- and N-terminally) of each Cdx/Cad member using NetPhos (Blom *et al.*, 1999). The serine residues in the conserved S-P motifs score between 0.975 and 0.998, strongly suggesting that this indeed represents a conserved phosphorylation

site. Finally, the linker region between the hexapeptide and the homeodomain contains the amino acid sequence g-K-T-R-T in all Cdx/Cad members, directly N-terminal to the homeodomain (Fig. 1B), defining a third caudal specific motif in addition to the homeodomain and the hexapeptide. The N-terminal arm of the homeodomain has been implicated in the establishment of target specificity of Hox proteins by functional studies in *Drosophila* (Kuziora and McGinnis, 1989; Gibson *et al.*, 1990) and X-ray diffraction and NMR studies (Jabet *et al.*, 1999; Passner *et al.*, 1999; Piper *et al.*, 1999).
The remarkably high sequence conservation between the Cdx family members suggests that amino acids in the linker region between homeodomain and hexapeptide contribute to target specificity. Taken together these results show that the Cdx family members contain a highly conserved hexapeptide that shares features with the first, third and fourth Hox paralog groups, but is Cdx specific. In addition, the strict sequence conservation in the linker region directly N-terminal to the homeodomain of Cdx proteins suggests a role for this region, and because it is located between the homeodomain and the Pbx interaction domain it is tempting to speculate about a role for the linker region in target specificity. Interestingly, Gsh family members do not have a hexapeptide in contrast to its closest Hox relatives (Fig 3A). *Pdx1* contains a hexapeptide, most similar in sequence to that of Hox paralog group 3 members, also the closest Hox group based on homeodomain sequence, further strengthening the suggested common ProtoHox origin of the Pdx1 and Hox paralog group 3 (Kourakis and Martindale, 2000). A scenario for successive duplication of genes and clusters has been proposed to account for the genomic organisation of the Antp-class of homeodomain proteins (Fig. 3) (Pollard and Holland, 2000). Accordingly, all four ProtoHox genes likely contained a hexapeptide. Interestingly, a hexapeptide sequence has been found in the Engrailed genes, raising the intriguing possibility that the hexapeptide origin could be even more widespread. To further investigate this we surveyed for putative hexapeptide sequences in all the Antp related homeodomain proteins.
In humans, the genes HB9, En2, and Gbx1 are linked, the same holds true for En1 and Gbx2 (Fig. 3A). These genes most likely arose by duplication of the so-called EHGbox array (Fig. 3B). The Engrailed genes have been shown to contain a hexapeptide sequence (Peltenburg and Murre, 1996). Gbx1 and Gbx2 do not contain a hexapeptide sequence, but interestingly, HB9 does. The hexapeptide and its flanking sequences are conserved between chicken, *Fugu rubripes*, human, mouse, and *Xenopus laevis* HB9 (Fig. 1C). The other cluster proposed to have arisen from the ArcheHox cluster is the ProtoNKL cluster (Fig. 3B). We found that Msx, Nkx, Hmx, Emx, and Tlx homologs

contain a putative hexapeptide sequence (Fig. 1D and 3A). Under the assumption of divergence of the hexapeptide, as opposed to independent acquisition, this suggests that an ancient form of the hexapeptide was present in all of the ArcheHox cluster members, and therefore presumably in the ancestor of the Antp-related proteins. Interestingly, the TALE-class of homeodomain co-factors has been shown to be very ancient, as a TALE-class factor was present in the common ancestor of plants, fungi, and animals (Bürglin, 1997). Our findings suggest that hexapeptide mediated interactions between Antp superfamily members and TALE-class co-factors appeared early during evolution.

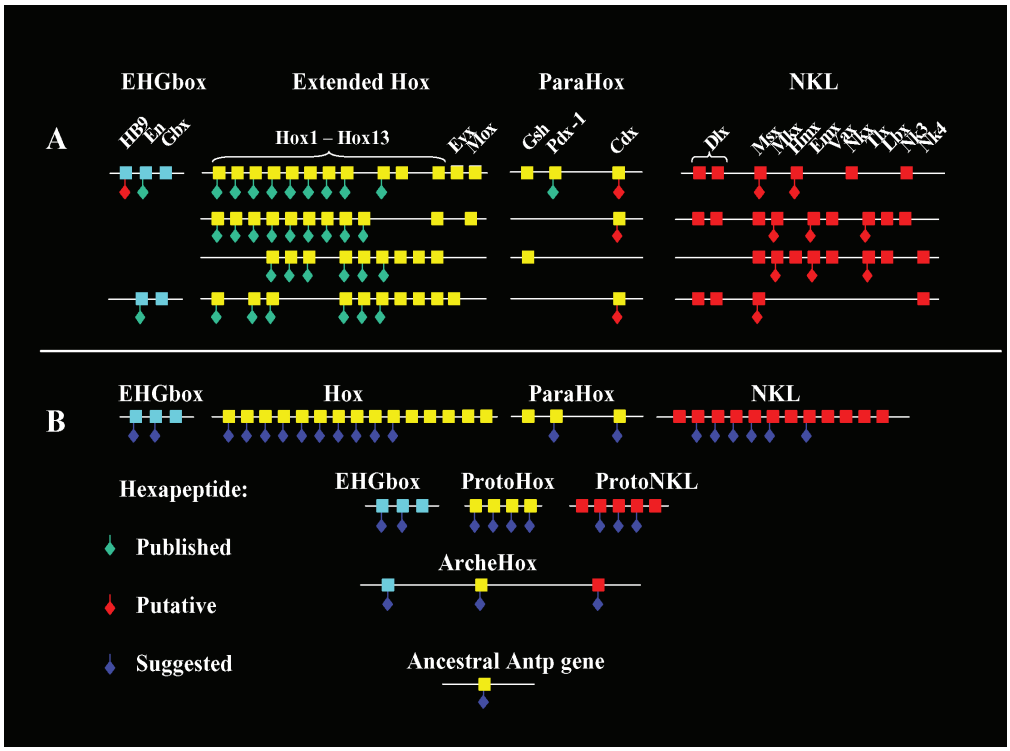

Figure 3. Distribution and proposed evolution of hexapeptide sequences in the Antp-class of homeobox genes. The schematic depiction of the clustral organisation is adapted from Pollard and Holland, 2000. **(A)** Overview of EHGbox, Extended Hox, ParaHox and NKL clusters. Green diamonds depict previously described hexapeptide sequences while red diamonds depict putative hexapeptides. **(B)** Suggested history of the hexapeptide, suggested hexapeptides are depicted by blue diamonds.

References

van den Akker, E., Forlani, S., Chawengsaksophak, K., de Graaff, W., Beck, F., Meyer, B. I., and Deschamps, J., (2002), Cdx1 and Cdx2 have overlapping functions in anteroposterior patterning and posterior axis elongation. Development 129: 2181-2193.

Blom, N., Gammeltoft, S., and Brunak, S., (1999), Sequence- and Structure-Based Prediction of Eukaryotic Protein Phosphorylation Sites. J. Mol. Biol. 294: 1351-1362.

Brooke, N. M., Garcia-Fernàndez, J., and Holland, P. W. H., 1998, The ParaHox gene cluster is an evolutionary sister of the Hox gene cluster. Nature vol. 392: 920-922.

Bürglin T. R., (1997), Analysis of TALE superclass homeobox genes (Meis, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals. Nucleic Acid Res. 25: 4173-4180.

Coulier, F., Burtey, S., Chaffanet, M., and Birnbaum, D., 2000, Ancestrally-duplicated paraHOX gene clusters in humans. Int. J. Oncol. 17(3): 439-444.

Di Rocco, G., Mavilio, F., and Zappavigna, V., (1997), Functional dissection of a transcriptional active, target-specific Hox-Pbx complex. EMBO J. 16:3644-3654.

Gibson, G., Schier, A., LeMotte, P., and Gehring, W. J., (1990), The specificity of Sex combs reduced and Antennapedia are defined by a distinct portion of each protein that includes the homeodomain. Cell 62: 1087-1103.

Goudet, G., Delhalle, S., Biemar, F., Martial, J. A., Peers, B., (1999), Functional and cooperative interactions between the homeodomain PDX1, Pbx, and Prep1 factors on the somatostatin promoter. J. Biol. Chem. 274: 4067-4073.

Holland, P. W., 2001, Beyond the Hox: how widespread is homeobox gene clustering? J. Anat. 199: 13-23.

Jabet, C., Gitti, R., Summers, M. F., and Wolberger, C., (1999), NMR Studies of the Pbx1 TALE Homeodomain Protein Free in Solution and

Bound to DNA: Proposal for a Mechanism of HoxB1-Pbx1-DNA Complex Assembly. J. Mol. Biol. 291: 521-530.

Kim, S. K., Selleri, L., Lee, J. S., Zhang, A. Y., Gu, X., Jacobs, Y., Cleary, M. L., 2002, Pbx1 inactivation disrupts pancreas development and in Ipf1-deficient mice promotes diabetes mellitus. Nat. Genet. 30(4): 430-435.

Knoepfler, P. S., Lu, Q., and Kamps, M. P., (1996), Pbx1-Hox heterodimers bind DNA on inseparable half-sites that permit intrinsic DNA binding specificity of the Hox partner at nucleotides 3' to a TAAT motif. Nucl. Acid Res. 24: 2288-2294.

Kourakis, M. J. and Martindale, M. Q., (2000), Combined-Method Phylogenetic Analysis of Hox and ParaHox Genes of the Metazoa. Mol. Dev. Evol. 288: 175-191.

Kuziora, M. A., and McGinnis, W., (1989), A homeodomain substitution changes the regulatory specificity of the deformed protein in Drosophila embryos. Cell 59: 563-571.

Mann, R. S. and Chan, S. K., (1996), Extra specificity from extradenticle: the partnership between HOX and PBX/EXD homeodomain proteins. Trends Genet. 12: 258-262.

Morgan, R., In der Rieden, P. M. J., Hooiveld, M. H. W., and Durston, A. J., (2000), Identifying HOX paralog groups by the PBX-binding region. Trends Genet. 16: 66-67.

Passner, J. M., Ryoo, H. D., Shen, L., Mann, R. S., and Aggarwal, A. K., (1999), Structure of a DNA-bound Ultrabithorax-Extradenticle homeodomain complex. Nature 397: 714-718.

Peltenburg, L. T. C. and Murre, C., (1996), Engrailed and Hox homeodomain proteins contain a related Pbx interaction motif that recognizes a common structure present in Pbx. EMBO J. 15: 3385-3393.

Piper, D. E., Batchelor, A. H., Chang, C. P., Cleary, M. L., Wolberger, C., (1999), Structure of a HoxB1-Pbx1 Heterodimer Bound to DNA: Role of the Hexapeptide and a Fourth Homeodomain Helix in Complex Formation. Cell 96: 587-597.

Pollard, S. L., and Holland, P. W. H., 2000, Evidence for 14 homeobox gene clusters in human genome ancestry. Curr. Biol. 10: 1059-1062.

Pöpperl, H., Bienz, M., Struder, M., Chan, S. K., Aparicio, S., Brenner, S., Mann, R. S., and Krumlauf, R., (1995), Segmental expression of *Hoxb1* is Controlled by a Highly conserved Autoregulatory Loop Dependent upon *exd*/*pbx*. Cell 81: 1031-1042.

Yanze, N., Spring, J., Schmidli, C., and Schmid, V., (2001), Conservation of *Hox/ParaHox*-Related Genes in the early Development of a Cnidarian. Dev. Biol. 236: 89-98.

Chapter 7

The small GTPase *Rap1* is an immediate downstream target for *Hoxb4* transcriptional regulation

Ali S. Morsi El-Kadi[a], Paul In der Rieden[b], Antony Durston[b], Richard Morgan[a],

[a] Department of Anatomy and Developmental Biology, St. George's Hospital Medical School, Cranmer Terrace, London SW17 0RE, UK.
[b] Hubrecht Laboratorium, Nederlands Instituut voor Ontwikkelingsbiologie, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands.

Reprinted from Mech. Dev. (2002) 113, 135-139

ELSEVIER

Mechanisms of Development 113 (2002) 131–139

www.elsevier.com/locate/modo

The small GTPase *Rap1* is an immediate downstream target for *Hoxb4* transcriptional regulation

Ali S. Morsi El-Kadi[a], Paul In der Rieden[b], Antony Durston[b], Richard Morgan[a,*]

[a]*Department of Anatomy and Developmental Biology, St. George's Hospital Medical School, Cranmer Terrace, London SW17 0RE, UK*
[b]*Hubrecht Laboratory, Netherlands Institute for Developmental Biology, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands*

Received 12 December 2001; received in revised form 7 February 2002; accepted 14 February 2002

Abstract

The Hox genes are a family of homeodomain-containing transcription factors which determine anteroposterior identity early on in development. Although a lot is now known about their regulation and function, very little is known of their effector (downstream target) genes. Here we show that the small GTPase *Rap1* is a direct, negatively regulated target of *Hoxb4* and is excluded from *Hoxb4* expressing cells.

Keywords: *Hoxb4*; *Rap1*; Early development; Small GTPase; Transcriptional control

1. Introduction

A number of molecular components employed in patterning the early embryo have been very highly conserved through evolution, and the *Hox* genes are a particularly noticeable example. They encode a family of homeodomain transcription factors that are expressed along the anterior to posterior (anteroposterior) axis from the gastrula stage onwards, in a spatial and temporal order that closely reflects their relative position within the chromosome. Their expression defines the anterior–posterior character of those cells which express them, and consequently their deletion or misexpression can cause the transformation of one part of the embryo into another (reviewed by Carroll, 1995; Gehring, 1998; Burke, 2000).

The control and function of *Hox* genes have been extensively studied in a wide range of species, and many of the molecular aspects of their regulation are now understood. However, far less is known about the actual molecular basis of *Hox* gene function. In vitro, most HOX proteins recognize the same four-base-pair consensus sequence that is actually repeated many times in the genome (Hayashi and Scott, 1990). Far greater binding specificity is achieved when HOX proteins bind as a complex with other proteins, including PBX (vertebrate homologues of *Drosophila* homeodomain-containing transcription factor extradenticle, Chang et al., 1996). Thus HOX proteins can recognize considerably more specific sites in vivo, the identity of which have now been tentatively established (Chang et al., 1995, 1996; Ryoo and Mann, 1999; White et al., 2000). In order to understand the molecular basis for *Hox* gene function, we need to know which genes are the immediate targets of transcriptional activation or repression. However, to date very few of the downstream targets of *Hox* genes are known.

For this reason we have looked for genes that are directly regulated by the *Hox* gene *Hoxb4*. This is the vertebrate homologue of the *Drosophila deformed* gene which is expressed in and required for the correct specification of a number of cephalic segments. *Deformed* mutants lack maxillary and mandibular structures, the head having been transformed to thoracic like structures dorsally and deleted ventrally (Merrill et al., 1987). Studies in the mouse and in the frog (Xenopus) have revealed that *Hoxb4* is expressed in the hindbrain and spinal cord, with a sharp boundary in the hindbrain between rhombomeres 6 and 7 (Graham et al., 1988; Godsave et al., 1994). The homozygous null mutation of *Hoxb4* in the mouse results in a homeotic transformation of the second cervical vertebrae from axis to atlas, and defective morphogenesis of the sternum (Ramirez-Solis et al., 1993). Ectopic expression of *Hoxb4* in early Xenopus embryos results in the deletion of structures anterior to where *Hoxb4* is usually expressed (i.e. the forebrain, midbrain and hindbrain anterior to rhombomere 7 (Hooiveld et al., 1999)).

* Corresponding author. Tel.: +44-20-8725-5219; fax: +44-20-8725-3326.
E-mail address: r.morgan@sghms.ac.uk (R. Morgan).

0925-4773/02/$ - see front matter
PII: S0925-4773(02)00047-3

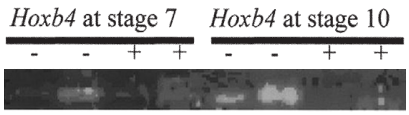

Fig. 1. Differential display identifies *XRap1* as a possible downstream target of *Hoxb4*. *Hoxb4/GR* (which confers dexamethasone dependence on *Hoxb4* activity) was injected into fertilized eggs and activated at either stage 7 (blastula) or stage 10 (gastrula). Total RNA was extracted at the neurula stage and randomly amplified. Identical but independent amplifications were performed to check for reproducibility. (+), Hoxb4 activated by dexamethasone at the stage indicated; (−), no dexamethasone added.

Here we report that the small GTPase *Rap1* is a direct target of *Hoxb4* regulation. *Hoxb4* represses *Rap1* expression in a manner that is independent of protein translation, and may bind to two putative HOXB4 protein binding sites located at the 3′ end of the *Rap1* gene in order to mediate this inhibition.

2. Results

2.1. XRap1 and Hoxb4 have complementary expression patterns in early development

In order to search for downstream targets of *Hoxb4*, we used a differential display technique to compare gene expression in embryos which had developed from eggs injected with *Hoxb4* RNA to that in untreated controls. Our attention was drawn to one transcript in particular because it was absent in *Hoxb4* injected embryos (Fig. 1). We therefore cloned and sequenced the corresponding cDNA from the untreated, control embryos. Conceptual translation of the partial open reading frame encoded in this clone gives a peptide which is 96% identical to the human RAP1 protein, a small GTPase (Pizon et al., 1988). The only differences in the amino acid sequence of the two proteins are 'conservative' changes (i.e. where the amino acids have very similar biochemical properties). We therefore conclude that the differentially expressed clone encodes the *Xenopus* homologue of the Human *Rap1* gene, and we refer to it here as X*Rap1* (accession number AY059389).

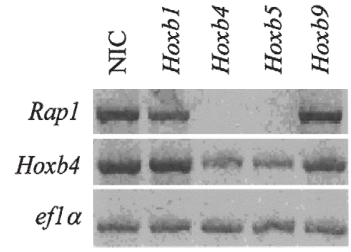

Fig. 2. RT–PCR analysis of RNA extracted from non-injected control ('NIC') or *Hox*-expressing embryos. Fertilized eggs were injected with either *Hoxb1*, *Hoxb4*, *Hoxb5* or *Hoxb9* mRNA (as shown above each lane). Total RNA was extracted at the neurula stage and examined for the expression of *XRap1*, *Hoxb4* or *ef1α* by RT–PCR (the latter is included as a loading control).

In order to confirm that *XRap1* is indeed repressed by *Hoxb4*, we injected fertilized eggs with *Hoxb4* mRNA and then examined the expression of *XRap1* later in development, at the neurula stage (Fig. 2). Using a similar approach, we also examined the affect that *Hoxb1*, *Hoxb5* and *Hoxb9* over-expression have on *XRap1* (Fig. 2). Both *Hoxb4* and *Hoxb5* result in a striking down-regulation of *XRap1*, whilst both *Hoxb1* (a more anteriorly expressed *Hox* gene) and *Hoxb9* (the most caudally expressed *Hoxb* gene) have no apparent affect on its expression.

In order to determine whether the apparent mutually exclusive relationship between X*Rap1* and *Hoxb4* is reflected in their expression pattern in the embryo, we used whole-mount in situ analysis to study their expression at a number of different developmental stages (Fig. 3). The expression of *XRap-1* begins early in gastrulation (Fig. 3A) in the dorsal ectoderm (the future neural plate). *Hoxb4* expression is detected slightly later in gastrulation in a more restricted pattern, being located at a more posterior and ventral position (Fig. 3B). *XRap1* remains confined to the neural tube during neurulation and is expressed throughout its length, with the exception of those cells that express *Hoxb4*. A sharp boundary between *XRap1* and *Hoxb4* expression domains becomes apparent as neurulation proceeds. By the tailbud stage the *XRap1* and *Hoxb4* domains of expression abut sharply both in the hindbrain just posterior to the otic vesicle (the boundary between rhombomeres 6 and 7; Fig. 3F), and in the head mesenchyme at the position of the fourth pharyngeal arch.

2.2. The repression of XRap-1 by Hoxb4 is direct and independent of protein synthesis

The preceding results indicate that *Hoxb4* represses *XRap-1* expression, but they do not provide any indication as to whether this repression is direct (i.e. independent of further translation), or indirect. In order to address this, we used a fusion between *Hoxb4* and the human glucocorticoid receptor (*Hoxb4*/GR, Hooiveld et al., 1999). The glucocorticoid receptor binds the heat shock protein HSP90, preventing it from entering the nucleus. This steric hindrance of nuclear entry is relieved by ligand binding, in this case the glucocorticoid analogue dexamethasone (DEX), which by itself has no discernible effects on *Xenopus* development (Gammill and Sive, 1997). Hence the *Hoxb4*/GR construct confers DEX dependence on the activity of *Hoxb4* (Hooiveld et al., 1999).

We injected fertilized eggs with *Hoxb4*/GR RNA and allowed them to develop to the mid-neurula stage. The embryos were then treated with dex in the presence or absence of cycloheximide (CHX), which blocks protein synthesis. We examined the expression of *XRap-1* and *Hoxb4* by reverse transcription–polymerase chain reaction (RT–PCR) of RNA subsequently extracted from these

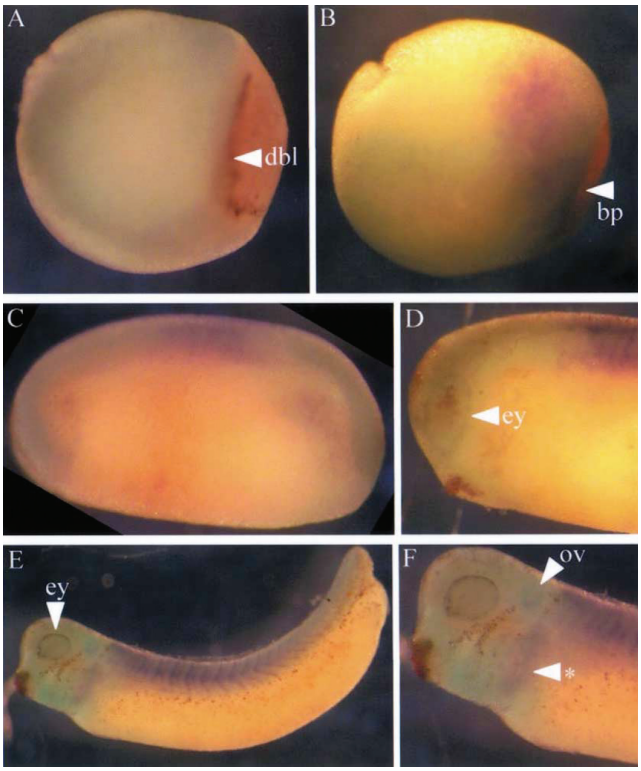

Fig. 3. Whole-mount in situ hybridization analysis of *Hoxb4* (purple stain) and X*Rap1* (light blue stain) expression. (A) A gastrula stage embryo (stage 10.5) with the blastopore on the right. *XRap1* staining is predominately in the dorsal ectoderm (which gives rise to the neural plate). No *Hoxb4* expression is observed at this stage. (B) A late gastrula/early neurula stage embryo (stage 12, lateral view with the dorsal side down). *Hoxb4* expression is apparent on the ventral/posterior side of the embryo and extends into the neural plate. X*Rap1* expression is confined to the neural plate. (C) Lateral view of a neurula (stage 20) embryo, anterior to left and dorsal side uppermost. *XRap1* expression persists in the whole of the neural tube with the exception of the *Hoxb4* expressing cells, roughly midway along the axis. (D) Lateral view of the anterior most (head) end of an early tailbud (stage 24) embryo. *XRap1* staining is in the anterior neural tube including the eye, and abuts the area of *Hoxb4* expression. (E) Lateral view of a tailbud (stage 28) embryo, anterior to left. (F) Enlargement of the anterior end of the embryo shown in (E). *XRap1* and *Hoxb4* expression domains abut closely in the hindbrain just posterior to the otic vesicle and in the head mesenchyme, where a stripe of *Hoxb4* expression (*) marks the fourth pharyngeal arch. bp, blastopore; dbl, dorsal blastopore lip; ey, eye; ov, otic vesicle.

embryos (Fig. 4). *Hoxb4* positively autoregulates its own expression by a direct mechanism (Hooiveld et al., 1999), thus activating the *Hoxb4*/GR construct should upregulate *Hoxb4* expression, even in the presence of cycloheximide, as indeed it does (Fig. 4).

The activation of *Hoxb4*/GR by DEX results in a strong down-regulation of X*Rap1*, indeed none can be detected when dex alone is added. There is also a very strong down-regulation of X*Rap1* when DEX and CHX are added together, although some X*Rap1* transcript is still detectable. This implies that the down-regulation of X*Rap1* by *Hoxb4* does involve a direct mechanism, at least in part. It should be noted though that this result is complicated slightly by the apparent upregulation of X*Rap1* by CHX (Fig. 4), which may counteract *Hoxb4* repression to some limited extent.

2.3. A HOXB4 consensus binding site present in the XRap1 gene can bind HOXB4 protein and mediate transcriptional repression by Hoxb4

The 3′ untranslated region (UTR) of the X*Rap1* gene contains two sites that are very similar to the consensus sequence determined for optimum binding of HOXB4 protein to DNA in vivo (White et al., 2000) (Fig. 5A). In order to determine whether these sites could mediate transcriptional repression by *Hoxb4*, we cloned them into a position immediately 3′ to a luciferase (luc) reporter gene, driven by a SV40 promoter (RP1B4+, Fig. 5A). This promoter drives expression of luc in *Xenopus* embryos (Fig. 5B; Etkin and Balcells, 1985). Co-injecting *Hoxb4* RNA with the RP1B4+ construct results in a significant down-regulation of luc activity. As a control, we used site

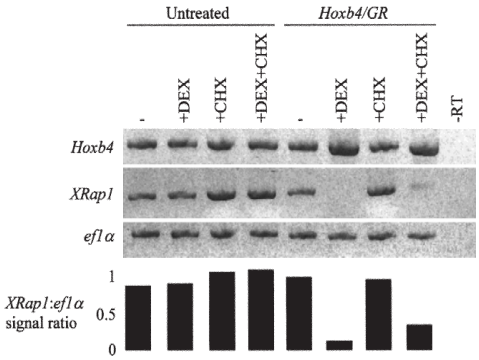

Fig. 4. RT–PCR analysis of RNA extracted from control ('untreated') or *Hoxb4*/GR expressing embryos. The embryos were treated with dexamethasone (DEX) and cycloheximide (CHX), either alone or in combination, as shown. *Efl α* is included as a loading control. −RT, PCR amplification without prior reverse transcription step. The *XRap1*:*Efl α* signal ratio is shown for each sample.

directed mutagenesis to alter the HOXB4 consensus binding sequence in RP1B4+. This second construct (RP1B4−) is not affected by *Hoxb4* co-injection. Additionally, we co-injected RP1B4+ with RNA transcribed from a deletion construct of *Hoxb4* that lacks its homeodomain (*Hoxb4*ΔHD) and has previously been shown not to upregulate *Hoxb4* expression (Hooiveld et al., 1999). *Hoxb4*ΔHD has no significant affect on RP1B4+ activity (Fig. 5B).

We also examined whether other Hox genes could also affect the expression of the reporter construct. To this end we also co-injected the RP1B4+ construct with *Hoxb1*, *Hoxb5* and *Hoxb9* mRNAs (Fig. 5C). Of these, only *Hoxb5* prevents the expression of the reporter gene, a finding that is in agreement with the data in Fig. 2.

Whilst the above data suggests putative HOXB4 binding site in the *XRap1* gene can mediate *Hoxb4* repression, they do prove that there is a direct interaction between them. We decided to examine this in vivo by using purified HOXB4 protein, linked to agarose beads via GST (glutathione-*S*-transferase). We incubated these beads with a PCR product derived from either the RP1B4+ or the RP1B4− construct,

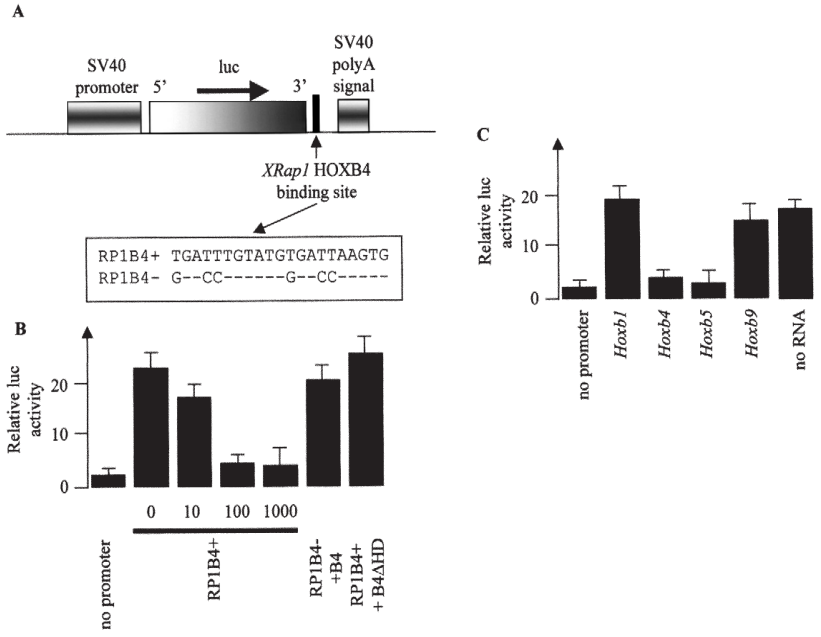

Fig. 5. The genomic region encoding the 3′ UTR of *XRap1* mRNA contains a putative HOXB4 binding site that can block the transcription of a luciferase (luc) reporter construct. (A) The reporter constructs were based on the pGL3 vector which contains the luc gene under the control of the ubiquitously active SV40 promoter. The putative HOXB4 binding site from *XRap1* was cloned immediately 3′ to the luc gene, as shown. The nucleotide sequence of the putative HOXB4 binding region (RP1B4+) and a non-binding variant used as a control (RP1B4−) are shown. (B) Co-injection of *Hoxb4* mRNA with the RP1B4+ reporter construct blocks its activity. The RP1B4+ reporter construct was injected into fertilized eggs together with *Hoxb4* mRNA (the amounts shown are in picograms (pg)), or with 1000 pg of a control *Hoxb4* mRNA lacking the homeodomain region ('RP1B4 + B4ΔHD'). The RP1B4− construct was co-injected with 1000 pg of *Hoxb4* RNA ('RP1B4− +B4'). 'No promoter', control luciferase construct lacking the SV40 promoter. (C) *Hoxb1*, *Hoxb4*, *Hoxb5* and *Hoxb9* RNAs (100 pg each) were also co-injected with RP1B4+ construct. The values represented in the figure are the means from three independent experiments, the error bars show the standard deviation. 'No RNA', RP1B4+ construct injected alone.

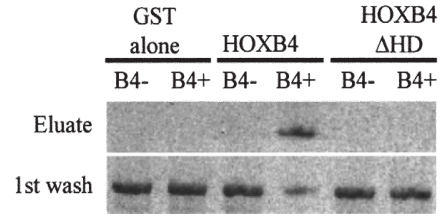

Fig. 6. HOXB4 protein can bind the genomic region encoding the 3′ UTR of *XRap1* mRNA. PCR amplified *XRap1* DNA, with (RP1B4+) or without (RP1B4−) the putative HOXB4 binding sequence, was incubated with HOXB4–GST (HOXB4) or GST alone on agarose beads. As a further control, a deletion construct of *Hoxb4* (lacking the homeodomain region), was also made as a GST construct and linked to beads (HOXB4ΔHD). The beads were washed to remove unbound probe and then probe that had bound was eluted. The amount of probe present in the first wash and eluate were determined by gel electrophoresis, as shown. B4−/B4+, RP1B4− or RP1B4+ DNA used in the incubation, respectively.

containing the putative HOXB4 binding site. After washing, only the RP1B4+-derived sequence remained bound to these beads (Fig. 6). As a control we also used beads with GST alone or with a deletion construct of *Hoxb4* lacking the homeodomain (HOXB4ΔHD). Neither of these proteins could bind the RP1B4+-derived sequence (Fig. 6).

2.4. Blocking the expression of the endogenous Hoxb4 gene results in a significant increase in XRap1 transcription

The translation of specific mRNAs in the early *Xenopus* embryo can be prevented by the injection of short DNA sequences that are chemically modified to prevent their degradation by endogenous nucleases ('morpholinos'). Morpholinos are designed to be complementary to the translation start site of the target mRNA, and have recently proved to be extremely effective at blocking the expression of specific target genes (Heasman et al., 2000; Ross et al., 2001). We designed a morpholino complementary to the translation start site of *Hoxb4* mRNA, together with a control morpholino of the same length and base composition but with a scrambled sequence. These were injected into fertilized eggs, and total RNA was extracted from neurula stage embryos that subsequently developed from them. We used this to examine the expression of *XRap1*, *Hoxb1*, *Hoxb4* and *Hoxb5* by RT–PCR (Fig. 7). The *Hoxb4* morpholino (B4morph) causes a significant down-regulation of the endogenous *Hoxb4* gene (Fig. 7, lane 2), presumably as a result of blocking its auto-regulation (Hooiveld et al., 1999). The control morpholino (conmorph) has no affect on *Hoxb4* expression. As an additional control on the specificity of the B4morph's specificity we co-injected B4morph with (human) *HOXB4* RNA. This prevents the B4morph mediated down-regulation of the endogenous *Hoxb4* gene (Fig. 7, lane 4). The B4morph, but not the conmorph, causes a significant increase in the expression of *XRap1*. Again, this effect is prevented by co-injection of *HOXB4* RNA (Fig. 7).

3. Discussion

3.1. Regulating XRap1

Here we have shown that the *Hox* gene *Hoxb4* directly represses the expression of a small GTPase, X*Rap1*. Furthermore *Hoxb4* and X*Rap1* have complementary expression patterns, with X*Rap1* being excluded from those cells which express *Hoxb4*. It is instead most strongly expressed in the anterior most part of the embryo which is fated to form the forebrain and midbrain.

Rap1 is a member of the *Ras* superfamily of small GTPases that cycle between a GTP bound (active) and a GDP bound (inactive) form (for a review see Bos et al., 2001). Their main function is to assemble and activate proteins at the cytoplasmic surface of membranes in response to specific stimuli. The most notable member of this family is *Ras* itself, with *Ras* mutations being present in 15% of all human tumours.

Rap1 antagonizes *Ras* signaling, probably by trapping one of its key effectors, the serine/threonine kinase *Raf*, in an inactive complex. Like Ras, *Rap1* is activated by guanine exchange factors (GEFs) that catalyse the exchange of GDP for GTP, and is inactivated by GTPase activating proteins

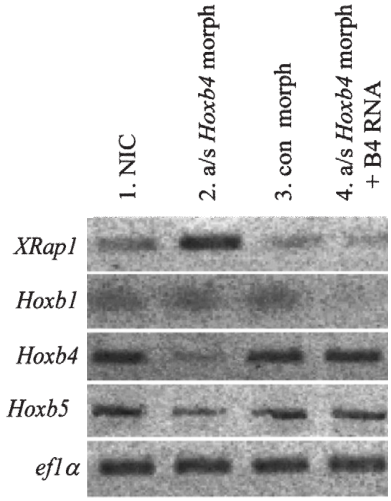

Fig. 7. Blocking *Hoxb4* translation results in the increased expression of *XRap1*. Fertilized eggs were injected with either an antisense *Hoxb4* morpholino (lane 2), a control (scrambled sequence) morpholino (lane 3), or both the antisense *Hoxb4* morpholino and *Hoxb4* mRNA (lane 4). The untreated, non-injected control is shown in lane 1. Total RNA was extracted from the embryos at the mid-neurula stage and analysed for the expression of *XRap1*, *Hoxb1*, *Hoxb4*, *Hoxb5* and *efl α* (a loading control), as shown.

(GAPs), that stimulate the otherwise very slow enzymatic activity of *Ras* family proteins. *Rap1*-interacting GEFs are activated either directly or indirectly by a number of intracellular messengers such as cAMP (de Rooij et al., 1998; Kawasaki et al., 1998a), Ca^{2+} and diacyl glycerol (DAG) (Kawasaki et al., 1998b; Yamashita et al., 2000; Ebinu et al., 1998). Far less is actually known about how *Rap1* is controlled at the transcriptional level, however. To our knowledge, its repression by *Hoxb4* is the first example of control being exerted at the transcriptional level during development.

3.2. Rap1 in development

Rap1 mediates a number of cellular processes that are involved in brain development. Most notably, it is essential for neurite outgrowth (York et al., 1998) and integrin-mediated cell adhesion (reviewed by Bos et al., 2001). Furthermore, mutations in the *Drosophila* homologue of *Rap1* disrupt the normal development of the eye (Karpilow et al., 1989). It is conceivable that there is a similar requirement for *Rap1* in vertebrate eye development, a possibility that is supported by our observation that X*Rap1* is expressed in the developing eye (Fig. 3) and, further, by a previous study showing that antibodies to β-integrin (a *Rap1* effector) block retinotectal projection in *Xenopus* (Stone and Sakaguchi, 1996).

Why is *XRap1* excluded from the more posterior regions of the embryo? Presumably *XRap1* activity may disrupt the normal development of the spinal cord by blocking specific developmental events, or by activating inappropriate ones. Hopefully this will be addressed by *XRap1* over-expression studies.

3.3. Transcriptional control of Rap1

As discussed above, *Rap1* is subject to numerous, stringent controls at the post-translational level. What is the likely significance of the very stringent transcriptional control we observe in early development? This is another question which is worthy of investigation. One possibility though is that, unlike the situation in the adult organism (or the cell lines from which they are derived), the complex post translational controls on *Rap1* activity have yet to be established. It may be then, that they only way to ensure the 'silence' of X*Rap1* in *Hoxb4* expressing cells is to prevent its transcription.

3.4. Hox downstream targets

Identifying the downstream targets of *Hox* genes is necessary if their function in development is to be understood in molecular terms. *XRap1* joins a very short list as to date very few *Hox* targets have been identified, and those which have tend to come from some what disparate systems. Furthermore, it is not always clear whether the identified targets are actually under the direct control of the *Hox* gene in question. Table 1 lists some of the *Hox* gene targets that have been identified in vertebrates (excluding the *Hox* genes themselves, which are subject to extensive auto- and cross-regulation (Gerard et al., 1996; Gould et al., 1997; Nonchev et al., 1997; Sharpe et al., 1998; Hooiveld et al., 1999)). The data is striking only for the lack of any apparent sequence or functional similarity between targets, and hints at an enormous complexity of downstream target control. In the light of the recent development of far more powerful techniques for detecting changes in gene expression, such as microarray analysis, it is slightly surprising that the list of *Hox* targets has grown so little. It is our hope that future studies will continue to address this fascinating problem.

4. Materials and methods

4.1. Differential display analyses

Fertilized *Xenopus* eggs were injected with 500 pg of *Hoxb4/GR* RNA. Dexamethasone was added at either stage 7 or 11, and the embryos were then allowed to develop until stage 17. At this point total RNA was extracted and

Table 1
Some *Hox* gene targets that have been identified in vertebrates

Hox gene	Transcription repressed (−) or activated (+)	Target gene	Target gene class	Ref.
Hoxb-1	+	COL5A2	Collagen	Penkov et al., 2000
Hoxb-3	+	TTF-1	Transcription factor	Guazzi et al., 1994
Hoxb-4	−	Xrap1	Small GTPase	This report
Hoxa-5	+	p53	Tumour suppressor	Raman et al., 2000
Hoxb-5	+	SPI3	Serine protease inhibitor	Safaei, 1997
Hoxb-7	+	bFGF	Secreted protein	Care et al., 1996
Hoxb-8/9	+	NCAM	Cell adhesion molecule	Jones et al., 1992
Hoxc-8	?	1(2)gl	Tumour suppressor	Tomotsune et al., 1993
Hoxa-9	−	Osteopontin	Secreted protein	Shi et al., 2001
Hoxa-10	+	p21	Kinase inhibitor	Bromleigh and Freedman, 2000
Hoxc-13	−	Keratin	Structural protein	Tkatchenko et al., 2001
Hoxa-9, a-10, b-9, c-9, d-10	+	Ren-1(c)	Renin	Pan et al., 2001

1μg was used to make cDNA by reverse transcription using a poly-deoxythymidine primer (T15). Two percent of this reaction was then randomly amplified by PCR using a single primer (5′-CAG ATT GGT GCT GGA TAT GC-3′), with two rounds of amplification at 94 °C for 30 s, 45 °C for 30 s and 72 °C for 60 s, and then 30 rounds of amplification at 94 °C for 30 s, 60 °C for 30 s and 72 °C for 60 s. The PCR products were resolved by electrophoresis on 2% agarose for 4 h at 200 V (4 °C) and visualized by ethidium bromide staining. Differentially displayed bands of interest were cut out the gel and the PCR products were extracted by Qiaquick PCR Purification kit (Qiagen) and eluted in 50 μl water. The purified PCR products were PCR-re-amplified and gel-purified if necessary, cloned into pGEM-T easy vector (Promega), and sequenced.

4.2. RNA extraction and RT–PCR

Total RNA was extracted from whole embryos using the QuickPrep Total RNA extraction kit (Amersham Pharmacia Biotech Inc.). 3 μg of RNA was used in subsequent reverse transcription reactions. This was mixed with a poly T_{15} oligo to 5 μg/ml and heated to 75 °C for 5 min. After cooling on ice, the following additional reagents were added; dNTPs to 0.4 mM, RNase OUT (Promega) to 1.6 units/μl, Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLRvT) RnaseH − point mutant (Promega) to 8 units/μl and the appropriate buffer (supplied by the manufacturer) to 1× concentration. The mixture was incubated for 1 h at 37 °C, heated to 70 °C for 2 min and cooled on ice.

PCR reactions were all performed in a total volume of 40 μl. For each we used 1 μl of the M-MLRvT reaction (as described above), 0.2 nmol of each primer and 20 μl of Redimix pre-mixed PCR components (Sigma). All reactions were cycled at 94 °C for 30 s, 60 °C for 30 s and 72 °C for 60 s. Thirty cycles were used for all primer sets except those for *ef1α*, for which 23 cycles were used. The primers used for X*Rap1* amplification were: forward – *XRAP1U*: 5′-GAT ATG CTG GGG GTG AAG CC-3′ and reverse – *XRAP1D*: 5′-CTA TTT AAT GTT CTT CAG CAG-3′.The sequences of the other primer pairs can be found on the internet at http://www.sghms.ac.uk/depts/anatomy/pages/richhmpg.htm/

4.3. Embryo culture and microinjection

These were performed as described previously (Sive et al., 2000).

4.4. Whole-mount in situ hybridization

X*Rap1* was cloned into vector pGEMT-easy (Promega), and this was linearized using *Hin*dIII. A fluorescein-labelled in situ probe was transcribed from this template using SP6 polymerase. A DIG-labelled *Hoxb4* probe was transcribed as previously described (Godsave et al., 1994). Probe purification and subsequent double probe in situ analysis were performed as described (Sive et al., 2000), using BCIP alone for *XRap1* detection and then NBT/BCIP to detect *Hoxb4*.

4.5. Luciferase reporter constructs

The putative HOXB4 binding sequences in the *XRap1* 3′ UTR region were cloned into the *Xba*I site of the pGL3 luciferase reporter construct (Promega), immediately 3′ to the luc reading frame (Fig. 5). In order to do this, the following oligos were synthesized: RP1B4 + U 5′-CTAGT GATTT GTATG TGATT AAGTG-3′; RP1B4 + D 5′-CTAGC ACTTA ATCAC ATACA AATCA-3′; RP1B4 − U 5′-CTAGG GACCT GTATG GGACC AAGTG-3′; RP1B4 − D 5′-CTAGC ACTTG GTCCC ATACA GGTCC-3′. Each of these four oligos were phosphorylated in separate reactions using polynucleotide kinase (PNK), using the protocol recommended by the manufacturer, and then the two (+) and (−) oligos were annealed by mixing half of each PNK reaction together, heating to 90 °C for 5 min and then cooling on ice. The annealed RP1B4+ and RP1B4− oligos were then ligated into pGL3 which had been restricted with *Xba*I, dephosphorylated using calf intestinal phosphatase (Promega), and purified using the Concert PCR purification system (Life Technologies). RP1B4+ and RP1B4− clones were selected that contained only one copy of the insert, and these were checked by sequencing. The chosen clones were then purified using the Plasmid Midi kit (Qiagen).

RP1B4− and RP1B4+ were injected into fertilized *Xenopus* eggs (100 pg in 5 nl), using the further refinements described by Mayor et al. (1993). Luciferase activity was measured as previously described (Morgan et al., 1999).

4.6. In vitro HOXB4 protein/Rap1 DNA interaction

A HOXB4/GST fusion protein was made by cloning the full-length *Hoxb4* reading frame into the pGEX-2TK vector (Amersham Pharmacia Biotech); further deletion constructs were then made from this base. The fusion proteins were produced and purified according to the manufacturers instructions (GST purification module, Amersham Pharmacia Biotech). *XRap1* DNA was added to the purified proteins (still attached to the agarose beads), and incubated for 5 min at room temperature in 50 mM NaCl, 2 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$ and 5 mM HEPES buffer (pH 7.4). Each of the DNAs tested were present at a final concentration of 1pg/ml, yeast tRNA was also added to a final concentration of 1 mg/ml.The beads were washed ten times in binding buffer (5 min each) and the bound RNA eluted by adding TE buffer and heating for 5 min at 50 °C.

4.7. Cycloheximide and dexamethasone treatments of Hoxb4/GR injected embryos

These were performed as described (Gammill and Sive, 1997). Embryos were incubated with cycloheximide for 30 min prior to the addition of dexamethasone. RNA was

extracted from embryos 2 h after dexamethasone treatment, by which point the untreated control embryos had reached stage 17.

Acknowledgements

RM wishes to thank St. George's Hospital Medical School for their kind support, and A. Morgan, T. King and S. Webb for their useful comments on the manuscript.

References

Bos, J.L., de Rooij, J., Reedquist, K.A., 2001. *Rap1* signalling: adhering to new models. Nat. Rev. Mol. Cell Biol. 2, 369–377.

Bromleigh, V.C., Freedman, L.P., 2000. p21 is a transcriptional target of HOXA10 in differentiating myelomonocytic cells. Genes Dev. 14, 2581–2586.

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

Care, A., Silvani, A., Meccia, E., Mattia, G., Stoppacciaro, A., Parmiani, G., Peschle, C., Colombo, M.P., 1996. HOXB7 constitutively activates basic fibroblast growth factor in melanomas. Mol. Cell Biol. 16, 4842–4851.

Carroll, S.B., 1995. Homeotic genes and the evolution of arthropods and cordates. Nature 376, 479–485.

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

Chang, C.P., Brocchieri, L., Shen, W.F., Largman, C., 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.

de Rooij, J., Zwartkruis, F.J., Verheijen, M.H., Cool, R.H., Nijman, S.M., Wittinghofer, A., Bos, J.L., 1998. Epac is a *Rap1* guanine-nucleotide-exchange factor directly activated by cyclic AMP. Nature 396, 474–477.

Ebinu, J.O., Bottorff, D.A., Chan, E.Y., Stang, S.L., Dunn, R.J., Stone, J.C., 1998. RasGRP, a Ras guanyl nucleotide- releasing protein with calcium- and diacylglycerol-binding motifs. Science 280, 1082–1086.

Etkin, L.D., Balcells, S., 1985. Transformed Xenopus embryos as a transient expression system to analyze gene expression at the midblastula transition. Dev. Biol. 108, 173–178.

Gammill, L.S., Sive, H., 1997. Identification of otx2 target genes and restrictions in ectodermal competence during Xenopus cement gland formation. Development 124, 471–481.

Gehring, W.J., 1998. Master Control Genes in Development and Evolution: The Homeobox Story, Yale University Press, New Haven/London.

Gerard, M., Chen, J.Y., Gronemeyer, H., Chambon, P., Duboule, D., Zakany, J., 1996. In vivo targeted mutagenesis of a regulatory element required for positioning the Hoxd-11 and Hoxd-10 expression boundaries. Genes Dev. 10, 2326–2334.

Godsave, S., Dekker, E.J., Holling, T., Pannese, M., Boncinelli, E., Durston, A., 1994. Expression patterns of Hoxb genes in the Xenopus embryo suggest roles in anteroposterior specification of the hindbrain and in dorsoventral patterning of the mesoderm. Dev. Biol. 166, 465–476.

Gould, A., Morrison, A., Sproat, G., White, R.A., Krumlauf, R., 1997. Positive cross-regulation and enhancer sharing: two mechanisms for specifying overlapping Hox expression patterns. Genes Dev. 11, 900–913.

Graham, A., Papalopulu, N., Lorimer, J., McVey, J.H., Tuddenham, E.G., Krumlauf, R., 1988. Characterization of a murine homeo box gene, Hox-2.6, related to the Drosophila Deformed gene. Genes Dev. 2, 1424–1438.

Guazzi, S., Lonigro, R., Pintonello, L., Boncinelli, E., Di Lauro, R., Mavilio, F., 1994. The thyroid transcription factor-1 gene is a candidate target for regulation by Hox proteins. EMBO J 13, 3339–3347.

Hayashi, S., Scott, M.P., 1990. What determines the specificity of action of Drosophila homeodomain proteins? Cell 63, 883–894.

Heasman, J., Kofron, M., Wylie, C., 2000. Beta-catenin signaling activity dissected in the early Xenopus embryo: a novel antisense approach. Dev. Biol. 222, 124–134.

Hooiveld, M.H., Morgan, R., in der Rieden, P., Houtzager, E., Pannese, M., Damen, K., Boncinelli, E., Durston, A.J., 1999. Novel interactions between vertebrate Hox genes. Int. J. Dev. Biol. 43 (7 Spec. No.), 665–674.

Jones, F.S., Prediger, E.A., Bittner, D.A., De Robertis, E.M., Edelmanm, G.M., 1992. Cell adhesion molecules as targets for HOX genes: neural cell adhesion molecule promoter activity is modulated by cotransfection with HOX-2.5 and -2.4. Proc. Natl. Acad. Sci. USA 89, 2086–2090.

Karpilow, J., Kolodkin, A., Bork, T., Venkatesh, T., 1989. Neuronal development in the Drosophila compound eye: rap gene function is required in photoreceptor cell R8 for ommatidial pattern formation. Genes Dev. 3, 1834–1844.

Kawasaki, H., Springett, G.M., Mochizuki, N., Toki, S., Nakaya, M., Matsuda, M., Housman, D.E., Graybiel, A.M., 1998a. A family of cAMP-binding proteins that directly activate *Rap1*. Science 282, 2275–2279.

Kawasaki, H., Springett, G.M., Toki, S., Canales, J.J., Harlan, P., Blumenstiel, J.P., Chen, E.J., Bany, I.A., Mochizuki, N., Ashbacher, A., Matsuda, M., Housman, D.E., Graybiel, A.M., 1998b. A Rap guanine nucleotide exchange factor enriched highly in the basal ganglia. Proc. Natl. Acad. Sci. USA 95, 13278–13283.

Mayor, R., Essex, L.J., Bennett, M.F., Sargent, M.G., 1993. Distinct elements of the xsna promoter are required for mesodermal and ectodermal expression. Development 119, 661–671.

Merrill, V.K., Turner, F.R., Kaufman, T.C., 1987. A genetic and developmental analysis of mutations in the Deformed locus in *Drosophila melanogaster*. Dev. Biol. 122, 379–395.

Morgan, R., Hooiveld, M.H., in der Reiden, P., Durston, A.J., 1999. A conserved 30 base pair element in the Wnt-5a promoter is sufficient both to drive its' early embryonic expression and to mediate its' repression by otx2. Mech. Dev. 85, 97–102.

Nonchev, S., Maconochie, M., Gould, A., Morrison, A., Krumlauf, R., 1997. Cross-regulatory interactions between Hox genes and the control of segmental expression in the vertebrate central nervous system. Cold Spring Harbor Symp. Quant. Biol. 62, 313–323.

Pan, L., Xie, Y., Black, T.A., Jones, C.A., Pruitt, S.C., Gross, K.W., 2001. An Abd-B class HOX.PBX recognition sequence is required for expression from the mouse Ren-1c gene. J. Biol. Chem. 276, 32489–32494.

Penkov, D., Tanaka, S., Di Rocco, G., Berthelsen, J., Blasi, F., Ramirez, F., 2000. Cooperative interactions between PBX, PREP, and HOX proteins modulate the activity of the alpha 2(V) collagen (COL5A2) promoter. J. Biol. Chem. 275, 16681–16689.

Pizon, V., Chardin, P., Lerosey, I., Olofsson, B., Tavitian, A., 1988. Human cDNAs *Rap1* and rap2 homologous to the Drosophila gene Dras3 encode proteins closely related to ras in the 'effector' region. Oncogene 3, 201–204.

Raman, V., Martensen, S.A., Reisman, D., Evron, E., Odenwald, W.F., Jaffee, E., Marks, J., Sukumar, S., 2000. Compromised HOXA5 function can limit p53 expression in human breast tumours. Nature 405, 974–978.

Ramirez-Solis, R., Zheng, H., Whiting, J., Krumlauf, R., Bradley, A., 1993. *Hoxb4* (Hox-2.6) mutant mice show homeotic transformation of a cervical vertebra and defects in the closure of the sternal rudiments. Cell 73, 279–294.

Ross, J.J., Shimmi, O., Vilmos, P., Petryk, A., Kim, H., Gaudenz, K., Hermanson, S., Ekker, S.C., O'Connor, M.B., Marsh, J.L., 2001.

Twisted gastrulation is a conserved extracellular BMP antagonist. Nature 410, 479–483.

Ryoo, H.D., Mann, R.S., 1999. The control of trunk Hox specificity and activity by Extradenticle. Genes Dev. 13, 1704–1716.

Safaei, R., 1997. A target of the HoxB5 gene from the mouse nervous system. Brain Res. Dev. Brain Res. 100, 5–12.

Sharpe, J., Nonchev, S., Gould, A., Whiting, J., Krumlauf, R., 1998. Selectivity, sharing and competitive interactions in the regulation of Hoxb genes. EMBO J. 17, 1788–1798.

Shi, X., Bai, S., Li, L., Cao, X., 2001. Hoxa-9 represses transforming growth factor-beta-induced osteopontin gene transcription. J. Biol. Chem. 276, 850–855.

Sive, H.L., Grainger, R.M., Harland, R.M., 2000. Early Development of *Xenopus laevis*: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Stone, K.E., Sakaguchi, D.S., 1996. Perturbation of the developing Xenopus retinotectal projection following injections of antibodies against beta1 integrin receptors and N-cadherin. Dev. Biol. 180, 297–310.

Tkatchenko, A.V., Visconti, R.P., Shang, L., Papenbrock, T., Pruett, N.D., Ito, T., Ogawa, M., Awgulewitsch, A., 2001. Overexpression of Hoxc13 in differentiating keratinocytes results in downregulation of a novel hair keratin gene cluster and alopecia. Development 128, 1547–1558.

Tomotsune, D., Shoji, H., Wakamatsu, Y., Kondoh, H., Takahashi, N., 1993. A mouse homologue of the Drosophila tumour-suppressor gene l(2)gl controlled by Hox-C8 in vivo. Nature 365, 69–72.

White, R.A., Aspland, S.E., Brookman, J.J., Clayton, L., Sproat, G., 2000. The design and analysis of a homeotic response element. Mech. Dev. 91, 217–226.

Yamashita, S., Mochizuki, N., Ohba, Y., Tobiume, M., Okada, Y., Sawa, H., Nagashima, K., Matsuda, M., 2000. CalDAG-GEFIII activation of Ras, R-ras, and *Rap1*. J. Biol. Chem. 275, 25488–25493.

York, R.D., Yao, H., Dillon, T., Ellig, C.L., Eckert, S.P., McCleskey, E.W., Stork, P.J., 1998. *Rap1* mediates sustained MAP kinase activation induced by nerve growth factor. Nature 392, 622–626.

Summary

The *Hox* genes form a subset of the homeobox containing genes. The homeobox encodes a DNA binding motif, called the homeodomain. In most animal species the *Hox* genes are organised in one or more clusters. The number of genes present in a cluster varies between animal species; the number of clusters in each species also varies. The Hox clusters are thought to have arisen by tandem duplication of a single gene, followed, in vertebrates, by duplication of the cluster itself. As a consequence, *Hox* genes occupying the same relative position along the 5' to 3' chromosomal coordinate, named paralogous genes, share more similarities in sequence, expression pattern and function than do adjacent *Hox* genes on the same chromosome. One of the conserved features of the *Hox* genes is that of spatial colinearity of *Hox* expression along the AP axis and other embryonic axes. This means that paralogs located in the most 3' position of a cluster are expressed at a more anterior position than genes located at more 5' clustral locations. Vertebrate embryos also display a second form of colinearity, temporal colinearity, whereby genes located more 3' in the cluster are expressed earlier than genes located more 5' in the cluster.

It has been firmly established that *Hox* expression boundaries along the AP and other embryonic axes are correlated with structural identities. Generating correct *Hox* expression patterns is thus clearly essential for correct AP axis patterning. The aim of investigations presented in this thesis has been to gain more insight into the processes controlling expression and function of *Hox* genes during anteroposterior patterning.

Recent work in our research group has shown that a temporally colinear expression sequence of *Hox* genes is already present in the marginal zone mesoderm of *Xenopus* gastrulae. Most, if not all, previous work concerning *Hox* colinearity has been focused on colinear *Hox* gene expression in the neurectoderm. In chapter two, we present data to show that *Xwnt8* is directly upstream of *Hoxd1* in marginal zone mesoderm. This is the first example of an initiator of expression of a 3' *Hox* gene in a vertebrate.

An upstream regulator of *Hox* gene expression in the neurectoderm of vertebrates is retinoic acid (RA) (and or its derivatives) Retinoids can act via the nuclear receptors of the RAR and RXR family. These receptors form heterodimers and bind Retinoic Acid Response Elements (RAREs) in the promoters of target genes. RAREs have been found in the regulatory sequences of a number of labial- and deformed group *Hox* genes. We took advantage of the availability of the genomic sequences of all four Hox

clusters of mouse and human, to search for conserved RAREs in the Hox clusters, and the results are reported in chapter three.
Most Hox proteins contain a second conserved domain, in addition to the homeodomain, the hexapeptide. This domain is needed for Hox/PBC member interaction, and this interaction leads to increased binding specificity and -affinity. Sequence analysis of Hox proteins demonstrates interspecies sequence conservation among paralog group members in the sequences flanking the hexapeptide, which is reported in chapter four.
Nuclear localisation of PBC family members is controlled by competing nuclear import and export signals. Interaction of Meis family members with PBC members shields the nuclear export signal of PBC proteins, resulting in a net influx into the nucleus, modifying the activity of Hox proteins present. In zebrafish hindbrain development, a synergistic relation between *Hoxb1*, *Pbx4*, and *Meis3* has been shown, and was argued to directly induce the expression of *Hoxb1*. Since recent discoveries have shown that *Hox* genes are expressed in a colinear sequence in marginal zone mesoderm it was investigated whether a *Xenopus* Meis homolog, *XMeis3*, cooperates with Hox function during gastrula stages. In chapter five, we report that *XMeis3* is necessary for mesodermal and ectodermal *Hox* expression, and the progression of gastrulation.
In chapter six, conservation of hexapeptide-flanking sequences of Pdx1 and Cdx proteins, present in a wide range of species, resembling the conservation found in Hox group 1 through 8 proteins (Chapter four) is reported. More generally the presence of a hexapeptide sequences and conservation of flanking sequences in all of the members of the Antp-class of homeodomain proteins was investigated, and found to be widely distributed, in addition, hexapeptide flanking sequences conservation was found.
Identifying the downstream targets of *Hox* genes is necessary if their function in development is to be understood in molecular terms. *XRap1* joins a very short list as to date very few Hox targets have been identified. In chapter seven, the small GTPase *XRap1* is reported as a direct target of *Hoxb4* regulation. *Hoxb4* represses *XRap1* expression in a manner that is independent of protein translation, and may bind to two putative HOXB4 protein-binding sites located at the 3' end of the *Rap1* gene in order to mediate this inhibition.

Samenvatting in het Nederlands

De *Hox* genen vormen een subset van de homeobox bevattende genen. De homeobox codeert voor een DNA bindend motief, homeodomein genaamd. In de meeste diersoorten zijn de *Hox* genen georganiseerd in één of meer clusters. Het aantal genen in een cluster varieert per soort, zo ook het aantal clusters. De meeste gewervelde dieren hebben vier clusters, die waarschijnlijk voortgekomen zijn uit stapsgewijze duplicaties van een enkel gen, gevolgd door clusterduplicatie. Als gevolg van die clusterduplicatie zijn de genen die op vergelijkbare locaties in verschillende clusters liggen meer verwant aan elkaar in sequentie, expressiepatroon en functie dan *Hox* genen die naast elkaar in hetzelfde cluster liggen. Deze *Hox* genen vormen zo dertien zogenoemde paraloge groepen. Één van de geconserveerde eigenschappen van *Hox* genen is ruimtelijke colineariteit, wat inhoudt dat genen gelegen aan de 3' kant van een cluster tot expressie komen in een meer anteriore positie op de anteroposteriore as, en andere embryonale assen, dan genen in een meer 5' positie. In embryo's van gewervelde dieren vertonen de *Hox* genen een tweede vorm van colineariteit, de temporele. Dit houdt in dat genen aan de 3' kant van een cluster eerder tot expressie komen dan meer 5' gelegen genen.
Expressiepatronen van *Hox* genen hebben een scherpe anteriore grens. Deze grenzen zijn verbonden met structurele identiteit. Dit maakt een strikte controle op expressie van *Hox* genen essentieel voor correcte specificatie langs de embryonale kop-staart as. Dit proefschrift behandelt een aantal aspecten van de controle op expressie en functie van *Hox* genen tijdens embryonale ontwikkeling.
Recentelijk is in onze werkgroep aangetoond dat in mesoderm een colineair *Hox* expressiepatroon voorkomt. In hoofdstuk twee wordt signaaltransductie van *Xwnt8* beschreven als directe activator van *Hoxd1* expressie in mesoderm, terwijl de expressie van *Hoxb4* niet afhankelijk blijkt van *Xwnt8*. Dit is het eerste voorbeeld van een directe initiator van expressie van een 3' *Hox* gen in mesoderm in een vertebraat embryo.
Een activator van *Hox* gen expressie in het neurectoderm van vertebraten is retinoine zuur. Retinoiden kunnen expressie van target genen beïnvloeden via nucleaire receptoren van de RAR en RXR familie. Deze receptoren vormen heterodimeren en binden specifieke sequenties, zogenoemde RARE's, in de promotoren van target genen. RARE's zijn gevonden in de regulatoire sequenties van een aantal Hox genen. In hoofdstuk drie is van de beschikbaarheid van genomische sequenties van de vier Hox clusters van de

muis en de mens gebruik gemaakt om te zoeken naar nog niet beschreven, geconserveerde RARE's in de Hox clusters.
Naast het homeodomein bevatten de meeste Hox eiwitten een tweede geconserveerd domein, de hexapeptide. Dit domein is verantwoordelijk voor interactie met co-factoren van de PBC familie. Deze interactie draagt zorg voor een verhoging van de bindingsspecificiteit en -affiniteit van Hox eiwitten voor target sequenties. Sequentie analyse van Hox eiwitten brengt naar voren dat een paraloge groep specificiteit bestaat voor de hexapeptide flankerende sequenties, dit is beschreven in hoofdstuk vier.
De nucleaire localisatie van PBC eiwitten wordt onder andere gereguleerd door MEIS eiwitten. Deze factoren zijn essentieel voor de ontwikkeling van de achterhersenen, en het is aangetoond dat, in zebravis embryo's, *Hoxb1*, *Pbx4* en *Meis3* de expressie van *Hoxb1* induceren. Zoals vermeld is recentelijk ook colineaire Hox expressie aangetoond in mesoderm. Hoofdstuk vijf beschrijft dat *XMeis3* noodzakelijk is voor mesodermale en ectodermale Hox expressie en voor de voortschrijding van gastrulatie.
De Hox genen vormen samen met de ParaHox-, de EHG-box- en de NKL cluster genen de Antp-familie van homeobox genen. Naast de Hox eiwitten bevatten de Pdx1 en Cdx (ParaHox), en de Engrailed (EHG-box) eiwitten een hexapeptide domein. In hoofdstuk zes wordt beschreven dat in meer leden van de Antp-familie een hexapeptide voorkomt, zo wijd verbreid zelfs dat het aannemelijk is dat een hexapeptide domein al aanwezig was in de gezamenlijke voorouder van de Antp-familie.
Voor het volledig begrijpen van de werking van Hox eiwitten op moleculair niveau is het van belang om target genen en sequenties te vinden. Tot nu toe zijn echter slechts een beperkt aantal target genen bekent. Hoofdstuk zeven beschrijft *Xrap1* als een target gen voor *Hoxb4* regulatie. *Hoxb4* inhibeert de expressie van *Xrap1* onder omstandigheden waarbij eiwit synthese geblokkeerd is. De repressie wordt mogelijk gemedieerd door twee veronderstelde HOXB4 bindingssequenties, gesitueerd aan de 3' kant van het *Xrap1* gen.

Curriculum vitae

Paul Marius Josephus In der Rieden was born on May 20th 1974 in Emmen. He attended the Katholiek Drents College in Emmen, graduating in 1992. He then studied Molecular Sciences at the Wageningen Agricultural University. During his undergraduate work, he participated in two research projects at the Wageningen University, under supervision of Dr. H. W. J. Stroband at the Department of Experimental Animal Morphology and Cell Biology, and a second project under supervision of Prof. Dr. S. C. De Vries at the Department of Molecular Biology. He graduated in 1997. The work described in this thesis was started in 1997 and performed at the Hubrecht Laboratorium under supervision of Prof. Dr. A. J. Durston. Since September 2002 he is working as a postdoctoral researcher at the Hubrecht Laboratorium in the group of Prof. Dr. A. J. Durston.

Paul Marius Josephus In der Rieden werd op 20 mei 1974 geboren te Emmen. In 1992 behaalde hij zijn VWO diploma aan het Katholiek Drents College te Emmen. Vanaf 1992 studeerde hij moleculaire wetenschappen aan de Landbouwuniversiteit Wageningen. Hij liep stage bij Dr. H. W. J. Stroband bij (vakgroep Experimentele Diermorfologie en Celbiologie), en bij Prof. Dr. S. C. de Vries (vakgroep Moleculaire Biologie). Deze studie werd in 1997 afgerond. Vanaf 1997 voerde bij zijn promotie onderzoek uit, onder leiding van Prof. Dr. A. J. Durston, aan het Hubrecht Laboratorium te Utrecht. Sinds September 2002 is hij als postdoctoraal onderzoeker verbonden aan het Hubrecht Laboratorium, in de groep van Prof. Dr. A. J. Durston.

List of Publications

In der Rieden, P. M. J., Stevens, C., Samallo, J., Schipper, H., Te Kronnie, G., and Stroband, H. W. J., (1996), Expression of HOXB-3 in Carp (Cyprinus carpio) Embryos. Int. J. Dev. Biol. Supplement: 97-98.

Vroemen, C. W., Aarts, N., In der Rieden, P. M. J., Van Kammen, A., and De Vries, S. C., Identification of genes expressed during *Arabidopsis thaliana* embryogenesis using enhancer trap and gene trap *Ds*-transposons. NATO ASI Series, Vol. H 104. Cellular Integration of Signalling Pathways in Plant Development. Edited by F. Lo Schiavo, R. L. Last, G. Morelli, and N. V. Raikhel. Springer-Verlag Berlin Heidelberg 1998.

Hooiveld, M. H. W., Morgan, R., In der Rieden, P. M. J., Houtzager, E., Pannese, M., Damen, K., Boncinelli, E., and Durston, A. J., (1999), Novel interactions between vertebrate *Hox* genes. Int. J. Dev. Biol. 43: 665-674.

Morgan, R., Hooiveld, M. H. W., In der Rieden, P. M. J., and Durston, A. J., (1999), A conserved 30 base pair element in the Wnt-5a promoter is sufficient both to drive its' repression by otx2. Mech. Dev. 85: 97-102.

Morgan, R., In der Rieden, P. M. J., Hooiveld, M. H. W., and Durston, A. J., (2000), Identifying HOX paralog groups by the PBX-binding region. Trends Genet. 16: 66-67.

Morsi El-Kadi, A. S., In der Rieden, P. M. J., Durston, A. J., and Morgan, R., (2002), The small GTPase *Rap1* is an immediate downstream target for Hoxb4 transcriptional regulation. Mech. Dev. 113: 131-139.