**HOXD4 and regulation of the group 4 paralog genes**

Alastair Morrison1,†, Linda Ariza-McNaughton1, Alex Gould1, Mark Featherstone2 and Robb Krumlauf1,*

1Division of Developmental Neurobiology, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK
2McGill Cancer Center, Department of Oncology and Medicine, McGill University, 3655 Drummond Street, Montreal, Quebec, Canada H3G 1Y6

*Author for correspondence (e-mail: r-krumlauf@nimr.mrc.ac.uk)
†Current address: Vertebrate Development Laboratory, Imperial Cancer Research Fund, PO Box No 123, 44 Lincoln’s Inn Fields, London WC2A 3PX, UK

**SUMMARY**

From an evolutionary perspective, it is important to understand the degree of conservation of cis-regulatory mechanisms between paralogous Hox genes. In this study, we have used transgenic analysis of the human HOXD4 locus to identify one neural and two mesodermal 3’ enhancers that are capable of mediating the proper anterior limits of expression in the hindbrain and paraxial mesoderm (somites), respectively. In addition to directing expression in the central nervous system (CNS) up to the correct rhombomere 6/7 boundary in the hindbrain, the neural enhancer also mediates a three rhombomere anterior shift from this boundary in response to retinoic acid (RA), mimicking the endogenous Hox4 response. We have extended the transgenic analysis to Hoxa4 identifying mesodermal, neural and retinoid responsive components in the 3’ flanking region of that gene, which reflect aspects of endogenous Hox4 expression. Comparative analysis of the retinoid responses of Hoxd4, Hoxa4 and Hoxb4 reveals that, while they can be rapidly induced by RA, there is a window of competence for this response, which is different to that of more 3’ Hox genes. Mesodermal regulation involves multiple regions with overlapping or related activity and is complex, but with respect to neural regulation and response to RA, Hoxb4 and Hoxd4 appear to be more closely related to each other than Hoxa4. These results illustrate that much of the general positioning of 5’ and 3’ flanking regulatory regions has been conserved between three of the group 4 paralogs during vertebrate evolution, which most likely reflects the original positioning of regulatory regions in the ancestral Hox complex.

Key words: Hox gene, gene regulation, hindbrain, retinoic acid, transgenic mice, HOXD4, somite, evolution

**INTRODUCTION**

The vertebrate Hox paralogy group 4 contains four genes highly related to the *Drosophila Deformed (Dfd)* gene. While their expression patterns are similar, there are differences in both the anterior boundaries and relative timing of expression during embryogenesis (Gaunt et al., 1989; Geada et al., 1992; Hunt et al., 1991). For example Hoxa4, Hoxb4 and Hoxd4 have anterior limits in the hindbrain that map to the junction between rhombomeric (r) segments r6 and r7 (Hunt et al., 1991; Keynes and Krumlauf, 1994), but their boundaries in paraxial mesoderm are all offset by a single somite (Gaunt et al., 1989; Hunt and Krumlauf, 1992). Mutation of these paralogous genes in mice reveals a partial functional redundancy or compensation illustrating that they can work synergistically to pattern individual vertebral components (Horan et al., 1994, 1995a,b; Kostic and Capecchi, 1994; Ramirez-Solis et al., 1993). More recently, it has been shown that this redundancy can operate at the level of individual target genes (Gould et al., 1997). Hoxc4 is the most disparate member of the group with respect to both expression and function, as its anterior limit maps to the rostral spinal cord not the hindbrain (Geada et al., 1992) and loss-of-function mutations affect the patterning of more posterior vertebrae than the other group 4 genes (Boulet and Capecchi, 1996; Saegusa et al., 1996).

The vertebrate Hox clusters arose by the duplication of an ancestral complex (Duboule, 1994; Holland and Garcia-Fernandez, 1996; Krumlauf, 1992) and it is possible that each paralog may have conserved a subset or all of the *cis*-regulatory regions characteristic of their ancestral ortholog. These may have subsequently been modified in some of the paralogs to produce small variations in the timing or boundaries of expression. An alternative hypothesis is that, after duplication, each gene may have evolved control regions independently.

Analysis of the mouse Hoxa4 and Hoxb4 genes has previously shown that it is possible to reconstruct many aspects of their endogenous expression patterns using lacZ reporter genes in transgenic mice (Behringer et al., 1993; Whiting et al., 1991). Elements within the 5’ flanking region of Hoxa4 can direct proper mesodermal and neural expression (Behringer et al., 1993). Conversely, in the case of Hoxb4, an enhancer in the intron mediates the proper mesodermal and posterior neural domains of expression, while a second enhancer 3’ of the gene controls anterior neural expression (Gould et al., 1997; Whiting et al., 1991). The position, sequence and general function of these Hoxb4 enhancers is conserved between puffer fish, chick and mouse homologs (Aparicio et al., 1995; Morrison et al.,
1995). The differences between \textit{Hoxb4} and \textit{Hoxa4} in position and activities of regulatory regions suggested that, in general, control elements might not be conserved between paralogs.

Analysis of the 5’ flanking regions of a third group 4 member (\textit{Hoxd4}) has identified sequences conserved between mouse and human that mediate autoregulatory (Popperl and Feath-

erstone, 1992) and retinoic acid responses in tissue culture cells (Mavilio, 1993; Popperl and Feath-

erstone, 1993). In transgenic mice, the conserved sequence block containing the human retinoic acid response element (RARE) functions in vivo as a retinoic acid (RA)-responsive enhancer that directs neural-specific expression of a basal promoter \textit{lacZ} reporter, with a sharp boundary in the anterior spinal cord (Morrison et al., 1996). However, this is posterior to the r6/7 boundary of the endogenous \textit{Hoxd4} gene (Hunt et al., 1991). In contrast, the human autoregulatory region, in the absence of the RARE, is unable to direct reporter expression in transgenic mice (Morrison et al., 1996). While both the autoregulatory and RA responsive elements may contribute to the full expression pattern of the \textit{Hoxd4} gene, they do not set the appropriate anterior boundaries in either hindbrain or paraxial mesoderm, suggesting other regulatory regions must be required.

In this study, we use transgenic analysis to identify \textit{cis}-regulatory regions of the human \textit{HOXD4} gene that generate the correct anterior boundaries of expression in mesoderm and neural ectoderm. We also evaluated the \textit{Hoxd4} gene and compared the retinoid responses of \textit{Hoxd4}, \textit{Hoxb4} and \textit{Hoxa4}. The results provide a basis for comparing and contrasting the activities and positions of the regulatory regions from these three group 4 paralogs.

\section*{MATERIALS AND METHODS}

Transgene construction

Reporter constructs utilized the BGZ40 vector, containing the human \textit{\beta}-globin promoter linked to the \textit{lacZ} gene and SV40 polyadenylation signal (Yee and Rigby, 1993). Constructs #1-#15 contain fragments from the human \textit{HOXD4} clone pG3-4.4cos3.1 (Cianetti et al., 1990). Construct #1 contains a 4.2 kb \textit{EcoRI} fragment in the 5’-3’ orientation, blunt end cloned into the SalI site. Constructs #2, #3 and #5 were made by digesting construct #1 with NotI and either SalI or XbaI or BglIII, respectively. Construct #4 was made in two steps, with inverse PCR (Clackson et al., 1991) generating the 95 bp deletion of CR1 sequences in a 720 bp subclone first, using the oligonucleotide primers: direct-5’-CAGAGGCCACCATGGCTGCTGCGC-3’ and inverse-5’-CAGGTCAGGACCATGTTTGCTGC-3’. Construct #6 was generated by blunt ended cloning of the \textit{HOXD4} BglII-XbaI fragment in the 5’-3’ orientation. Constructs #7 and #8 were generated by digesting construct #6 with KpnI and either Mscl or DraI, respectively. Construct #9 was generated by digesting construct #1 with BglIII and SalI end filling, which removes the 1.65 kb internal fragment and self ligating. Construct #10 contained a 960 bp fragment of \textit{HOXD4} 3’ flanking sequence generated by PCR using the oligonucleotides (A) direct-5’-TATGCTTCAGACAGCTTACAAAGGG-3’ and (B) inverse-5’-ATTGCTTCAGACGCATTCGACTATGACAGG-3’. Constructs #11 carried in the 5’-3’ orientation a 1.5 kb XbaI-Apal \textit{HOXD4} fragment. Constructs #12-#15 were generated by inserting \textit{HOXD4} genomic fragments, generated by PCR into the XbaI site. The inserts for constructs #12-#14 were generated with the same 5’ primer (C) direct-5’-AATGATTTTGATTTTGTAGATCGAGG-3’ and the following 3’ primers respectively #12 (D) indirect-5’-CTTCTGAGCAGGATGACAGTGC-3’, #13 (E) indirect-5’-GAAG-GTCCAATGCTCAGAGATTGGG-3’, and #14 (F) indirect-5’-AAAGACGTCTAGACTATGACAGG-3’.

\section*{RESULTS}

\subsection*{Identification of \textit{HOXD4} regulatory regions}

In order to identify enhancer elements of the \textit{HOXD4} gene responsible for generating the proper anterior limits of expression in the neural tube at the r6/7 junction and in paraxial mesoderm at the s5/6 boundary (Gaunt et al., 1989; Hunt et al., 1991), we tested regions from the human locus in transgenic mice. Previously, we showed that regions 5’ of the \textit{HOXD4}
First we tested a 4.2 kb EcoRI genomic fragment containing the \( \text{HOXD} 4 \) intron, exon two and 3.4 kb of sequence extending past the translational stop codon (construct #1; Fig. 1), linked to a \( \text{lacZ} \) reporter gene. In seven independent insertions at developmental time points from 9.0-11.0 days post coitum (dpc), this construct directed both neural and mesodermal reporter expression, which extended up to a sharp r6/7 anterior limit in the hindbrain and up to the level of the s5/6 boundary in the paraxial mesoderm (Fig. 2B,C,E). Typical of Hox genes, the expression of both the endogenous \( \text{Hoxd} 4 \) mRNA and the \( \text{lacZ} \) reporter in the most anterior domain of somitic expression (s6), was at a lower level than in more caudal somites. The anterior boundaries of expression directed by this 3′ regulatory region are the same as the endogenous \( \text{Hoxd} 4 \) gene (Fig. 2A,D) and, combined with the posterior neural enhancer located 5′ of \( \text{HOXD} 4 \) (Morrison et al., 1996), they reconstruct a major part of the \( \text{Hoxd} 4 \) pattern. However, these 5′ and 3′ regions do not completely reconstruct the endogenous \( \text{Hoxd} 4 \) pattern because there is an absence of reporter staining in lateral plate mesoderm and differences in limb expression. This suggests that other control regions may be missing or that the gene may need to be in the normal chromosomal context for proper regulation of these aspects of expression (compare Fig. 2A,B and D,E).

**Mapping a proximal 3′ mesodermal enhancer**

To further map the mesodermal and neural control regions, we generated a series of deletion constructs (#2-#15; Fig. 1A). First we tested a 2.36 kb EcoRI-StuI fragment (construct #2) containing the \( \text{HOXD} 4 \) intron, exon 2 and 1.14 kb of 3′ sequence. In nine transgenic founders, neural expression was specifically lost, but we consistently observed the same s5/6 mesodermal expression pattern seen with construct #1 (data not shown). A shorter 2 kb EcoRI-XbaI fragment (construct #3) gave an identical mesodermal-restricted pattern in three transgenic embryos (Fig. 2F,J).

Initially, we focused on the intron within construct #3, because in previous work on the \( \text{Hoxb} 4 \) paralog we identified a mesodermal enhancer located in the equivalent region (Morrison et al., 1995; Whiting et al., 1994). Furthermore, within the \( \text{Hoxb} 4 \) intron, there is a highly conserved region (CR1) present not only in the mouse, chick and pufferfish homologs but also in the introns of the paralogous \( \text{Hoxd} 4 \) and \( \text{Hoxa} 4 \) genes from several species (Aparicio et al., 1995; Cianetti et al., 1990; Morrison et al., 1995). In addition, a conserved block related to CR1 is found within the introns of the \( \text{Hoxa} 7 \) (Haerry and Gehring, 1996) and \( \text{Hoxd} 9 \) (Zappavigna et al., 1991) genes. In all cases, the core of this block contains three consensus homeodomain binding sites, suggesting it may function as a Hox target in vertebrates (Haerry and Gehring, 1996). To examine the potential role of CR1 in the \( \text{HOXD} 4 \) enhancer (construct #3), we created a specific 95 bp deletion of CR1 (construct #4), but observed no change in the expression pattern (n=3; data not shown) compared with wild-type construct #3 or with other deletion variants (constructs #5-8; Fig. 2). Therefore, despite the extensive degree of evolutionary conservation, this result shows that CR1 is not necessary in this transgenic assay for the regulatory activity of the proximal \( \text{HOXD} 4 \) 3′ mesodermal enhancer.
sequences other than CR1 could be important for regulatory activity, so we tested a 710 bp EcoRI-BglII fragment containing only the HOXD4 intron and 165 bp of exon 2 sequences (construct #5). However, in a total of 18 founder embryos, we detected no lacZ gene expression in mesoderm, demonstrating that the HOXD4 intron alone is not sufficient for mesodermal regulation.

This suggested that the enhancer is located in the 1.2 kb BglII-XbaI fragment, which was confirmed by testing the activity of this region alone (construct #6). In four transgenic embryos, we detected mesodermal expression with a boundary at the somite 5/6 junction, identical to that obtained with constructs #2-#4 (data not shown). Transgenic embryos carrying two further deletions of the 1.2 kb enhancer (constructs #7 and #8) all had a s5/6 boundary of somitic expression identical to that of the full 1.2 kb enhancer and the endogenous Hoxd4 gene (Fig. 2G,H,K,L). Thus elements capable of mediating the correct pattern of HOXD4 expression in somites lie within a 3' 375 bp DraII-XbaI fragment.

A second distal mesodermal enhancer 3' of HOXD4

We made an internal deletion of the 1.2 kb mesodermal enhancer in the context of the initial 4.2 kb fragment (construct #9) to determine whether the remaining regions were sufficient to mediate the proper r6/7 neural expression. Surprisingly, the patterns of transgene expression were identical in the neural tube and mesoderm to those observed with construct #1 (data not shown). This suggested that there is a second mesodermal/somite enhancer within these sequences capable of directing the proper anterior boundary of expression. Since we found no regulatory activity within the intron alone (construct #5), we examined the remaining 3' region. A 960 bp fragment generated by PCR (construct #10), directed a consistent pattern of transgene expression in the paraxial mesoderm, with an anterior boundary at the s5/6 junction (Fig. 2I,M). However, we detected no expression in the hindbrain or anterior spinal cord of any of the embryos. This implies that somitic expression of HOXD4 is generated by at least two independent regulatory regions each capable of directing mesodermal expression up to the s5/6 boundary. Sequence comparisons between these two enhancers revealed no extended blocks of identity and a search of the Eukaryotic Transcription Factor Database (TFD) (Ghosh, 1993) for potential known cis-acting binding sites was not informative with respect to candidate regulators.

Interestingly, in several of the embryos with construct #10, we observed staining in the limb bud, lateral plate mesoderm and tail bud (Fig. 2I) that were not seen with construct #1 (Fig. 2B,E), suggesting that negative regulatory influences may have been deleted in this smaller construct. Since the lateral palate and limb domains correspond to the specific sites where construct #1 transgene expression did not recreate the endogenous pattern, it is possible that in the normal chromosomal context these negative influences do not occur and we are not missing critical cis-elements needed for positive activation of the gene in these tissues.
Mapping the HOXD4 neural enhancer

The results above implied that the components involved in regulating neural expression might be located in the 5' portion of the 1.9 kb Stul-EcoRI fragment (Fig. 1). This was confirmed using three different overlapping subfragments of this region (constructs #11-#13) (data not shown and Fig. 3A,E). In all cases, the anterior limit of expression extended up to a sharp r6/7 boundary with reporter staining gradually declining caudally. To further define the position of the r6/7 enhancer in the minimal 700 bp region (construct #13), we generated two partially overlapping subfragments. In five transgenic embryos carrying the 5' 405 bp (construct #14), reporter activity was never detected (data not shown). In contrast a 3' 468 bp fragment (construct #15) mediated a consistent pattern of neural-restricted expression (Fig. 3B,F) in six transgenic founder. However, the anterior limit of expression did not extend up to the proper r6/7 boundary, but mapped caudally to the hindbrain/spinal cord region. These results show that multiple elements in the HOXD4 enhancer are required for setting the proper r6/7 neural expression domain. Elements in the 3' part of the enhancer are able to direct neural-restricted expression, while elements in the 5' region are required to set the correct anterior limit.

The Hoxd4 regulatory elements are conserved between mouse and human

We wished to determine whether the position of 3' regulatory regions were also present in the mouse gene. By sequencing in the 3' flanking region of the mouse Hoxd4 locus, we identified a region corresponding to the minimal 700 bp human neural enhancer within a 1.6 kb XbaI fragment. In three founder embryos, this fragment (construct #16) mediated neural-restricted staining up to the r6/7 boundary (Fig. 3C,G). This pattern was indistinguishable from that obtained with the human HOXD4 neural enhancer demonstrating that the function and position of the regulatory region was conserved. We have tried to utilize sequence comparisons between these two elements to identify functionally conserved blocks, but the degree of identity over the entire 700 bp is so extensive (85%) that this approach has been uninformative.

The human HOXD4 neural enhancer was flanked by two mesodermal control elements all contained in a 2.6 kb Dral-EcoRI fragment (Fig. 1). Therefore, we used a 3.6 kb HindIII fragment (construct #17), which extended both 5' and 3' of the mouse neural enhancer, to assay for the presence of similar mesodermal control regions. In addition to the expected r6/7 neural domain, three of six transgenic embryos displayed strong mesodermal expression with the same s5/6 anterior limit as the endogenous Hoxd4 gene (Fig. 3D). The three remaining embryos showed the r6/7 neural expression but had either weak or patchy mesodermal expression where it was hard to assign a distinct anterior boundary (Fig. 3H). We assume these differences are due to differential influences exerted by the sites of integration. Therefore, as in the human gene, this 3.6 kb fragment contains at least one regulatory region flanking the neural enhancer capable of mediating the proper pattern and boundary of somite expression, suggesting that the relative positions of the mesodermal and neural enhancers are conserved between these species.

The 3' HOXD4 neural enhancer responds to retinoids

The RARE located in the 5' region of the HOXD4 locus has been shown to mediate a response to ectopic retinoic acid (RA) in transgenic embryos (Morrison et al., 1996), and we wanted to determine if it is involved in the overall response of the HOXD4 gene to RA through interaction with other 3' control elements such as the neural enhancer defined here. As a basis for comparison, we first characterised the RA response of the endogenous Hoxd4 gene in utero. Using conditions known to induce Hoxb1 (Conlon and Rossant, 1992; Marshall et al., 1992), exposure to RA at 7.5-7.75 dpc does not induce ectopic Hoxd4 expression (Fig. 5I) even though the gene is expression at that time in the primitive streak. However, there is a rapid and dynamic induction of Hoxd4 in the anterior hindbrain in response to RA treatments at 9.5 dpc, when Hoxb1 is no longer induced by RA. The response can be seen within 6 hours, when ectopic expression is expanded anteriorly over three rhombomeres (r4-r6) but, after 12-24 hours, there is a gradual reduction in r5 (Fig. 5A,E). Under these conditions, we did not observe any change in the mesodermal pattern of Hoxd4 gene expression following exposure to RA.

The RA response of the endogenous gene is significantly different to that seen with the 5' RARE elements directing reporter expression (Morrison et al., 1996) and, since it occurs largely within the hindbrain, we wished to examine whether the 3' neural enhancer was involved. We exposed transgenic founder embryos carrying either construct #16 (the mouse neural element) or construct #11 (the human neural element) to RA at 7.5-7.75 dpc or 9.5 dpc, and assayed for lacZ activity...
at 10.5 dpc. Treatments at 7.75 dpc did not alter reporter expression. However, in both cases with treatments at 9.5 dpc, we detected an anterior shift in transgene expression virtually identical to that observed with the endogenous gene, where ectopic staining was seen in r4 and r6 after 24 hours (Fig. 5B,C,F,G). Weak r5 staining was also seen in some cases (data not shown). This demonstrates that the 3′ neural enhancer is itself capable of mediating a retinoid response, and is not strictly dependent upon the 5′ RARE elements. Within the 3′ enhancer, there are no typical consensus RAREs with a spacing of 2 or 5 base pairs between the direct repeats, but there are some atypical direct repeats with a spacing of 0 and 4 base pairs and two widely dispersed consensus half sites. Hence it is not clear if the RA response is direct. The nature of the RA response mediated by the 3′ enhancer and the endogenous gene are so similar, this element is likely to be a major component involved in regulating the ectopic response in the hindbrain. Thus the RA response of the HOXD4 gene involves inputs from multiple control elements and they appear to have a common window of competence to respond, which opens between 8.5 and 9.5 dpc.

3′ regulatory elements in the mouse Hoxa4 gene

Regulatory elements that can mediate Hoxa4-like patterns of expression lie in sequences 5′ to this gene (Behringer et al., 1993). However, we were interested in testing whether additional elements capable of directing Hoxa4 expression lay in 3′ regions by analogy to Hoxb4 (Whiting et al., 1991) and the HOXD4/d4 results above. Therefore, a ~3.3 kb XbaI-ApaI 3′ fragment from Hoxa4 (Fig. 1C; construct #18) was tested. In transgenic embryos, we consistently observed neural and mesodermal expression with sharp anterior boundaries (data not shown) and we generated a stable line (AL34) to investigate this pattern in more detail (Fig. 4A).

In transgenic embryos from the AL34 line and transient founders, there was staining in lateral plate and paraxial mesoderm, with the same anterior somite boundary at s7/8 observed with the endogenous Hoxa4 gene (Fig. 4A,G). The mesodermal expression of this transgene (construct #18) is similar to that previously observed with 5′ regulatory elements from Hoxa4 (Behringer et al., 1993). Therefore, like HOXD4, there are multiple overlapping mesodermal enhancer activities surrounding the Hoxa4 gene some of which lie in its 3′ flanking region.

The anterior boundary of neural expression with this Hoxa4 transgene appeared to be caudal to the r6/r7 junction (Fig. 4A-D). To illustrate this more clearly, we bred the AL34 line with another line (ML19) that carries a lacZ reporter gene under the control of an r3/r5 enhancer from the Hoxb-2 gene (Sham et al., 1993). This provides convenient internal reference points for identifying specific rhombomeric boundaries (Fig. 4B-D). From the lateral view of a 10.5 dpc embryo (Fig. 4B) and a flat mount of the hindbrain region (Fig. 4C), it is clear that the anterior limit of the Hoxa4 transgene lies more than a single rhombomere’s width from the r5 marker, in the vicinity of the r7/8 junction. Interestingly, in Hoxa4 whole-mount in situ analysis at 9.5-10.5 dpc, we consistently found that the anterior expression limit of the endogenous gene was also caudal to the r6/r7 boundary (Fig. 4G,H). Thus the endogenous Hoxa4 neural pattern of expression at 10.5 dpc closely resembles the transgene expression with construct #18.

At later stages between 11.5 and 12.5 dpc, endogenous Hoxa4 expression extends more anteriorly up to the r6/r7 junction, like Hoxd4 and Hoxb4, indicating that there is a difference in the time at which these genes reach their respective anterior limits (Gaunt et al., 1989; Hunt et al., 1991). Hence we examined expression in the AL34 line at 11.5-12.5 dpc to determine if it would shift more anteriorly, but staining remained caudal to the r6/r7 junction (Fig. 4D, and data not shown). This indicates that the 3′ Hoxa4 enhancer does not mimic the endogenous pattern at later stages.

A deletion of the Hoxa4 construct removing 1 kb of the DNA sequences that lie immediately adjacent to the gene (construct #19), also directed neural-restricted expression in a manner identical to that seen with the full enhancer (Fig. 4E,F). However, in these embryos, expression in mesoderm was specifically abolished. This demonstrates that separable neural and mesodermal control regions also operate in the 3′ flanking region of the Hoxa4 gene, which may be a common property of the group 4 paralogs.

The response of Hoxa4 to retinoids

To further our comparisons with HOXD4/Hoxd4, we have also analysed the RA response of Hoxa4. To check the endogenous RA response first, we performed whole-mount in situ hybridisation on 10.5 dpc embryos, which had been exposed to RA in utero at 9.5 dpc, as for Hoxd4. While we did observe an RA response for Hoxa4, this was quite different in nature to the Hoxd4 response. There was an increase in the strength of the signal, suggesting higher levels of Hoxa4 transcripts, and a small anterior shift in the expression boundary. This shift was from a region near the spinalcord/hindbrain boundary in control embryos (Fig. 4G,H), up to an anterior limit at the r6/r7 junction (Fig. 5K,L), but with no expression in more anterior rhombomeres. The treatment at 9.5 dpc had no detectable effect upon the mesodermal pattern of expression and, like Hoxd4, RA treatments at earlier times also failed to induce Hoxa4 expression (data not shown). Thus the temporal window in competence to respond to RA is conserved between Hoxd4 and Hoxa4 but the response itself is different.

Embryos carrying the Hoxa4 3′ neural enhancer (line AL34, construct #18) and the r3/r5 marker (ML19) were also treated with RA at 9.5 dpc and analysed for transgene induction after 24 hours. lacZ expression shifted more anteriorly to the r6/r7 boundary, but only in ventral regions (Fig. 5D,H). Thus there is an RA response mediated by the Hoxa4 neural element, but this enhancer was not sufficient to recreate the full induction observed for the endogenous gene.

Neural enhancers and the retinoid response of group 4 paralogs

The correlation between the 3′ neural enhancers and the ectopic response to RA for both the Hoxd4 and Hoxa4 genes suggests that this might be a general characteristic of group 4 paralogs. To explore this further, we examined the RA response of Hoxb4 and of the 3′ r6/r7 neural enhancer (region A), which we have previously identified in transgenic studies (Aparicio et al., 1995; Morrison et al., 1995; Whiting et al., 1991). We have used both whole-mount in situ hybridisation and immunohistochemistry to monitor the RA response of the endogenous Hoxb4 gene and obtained identical results with both methods. In the neural ectoderm, there is a strong and rapid RA response
at both 8.5 dpc and 9.5 dpc, but no induction with treatments at 7.5-7.75 dpc (Fig. 6A-C and data not shown). At 9.5 dpc, there is initially a three rhombomere shift extending the anterior boundary from r6/7 up to the r3/4 junction (Fig. 6B), and this is followed by a downregulation in r5 (Fig. 6C). Again we detected no changes in mesodermal expression after RA treatment at 9.5 dpc. This response is identical to that observed with Hoxd4. Exposure to RA at 8.5 dpc resulted in an anterior shift of one rhombomere up to the r5/6 boundary (data not shown).

To examine the response of the 3' Hoxb4 neural enhancer, we applied the same RA treatment regime to transgenic embryos carrying the lacZ reporter under region A control (construct #9 in Whiting et al 1991). Like the endogenous gene, the transgene displayed the same temporal windows of response, giving a one rhombomere shift at 8.5 dpc and a three rhombomere anterior shift with dynamic downregulation in r5 at 9.5 dpc (Fig. 6D-I; and data not shown). Furthermore, to confirm that the transgene and endogenous patterns were coincident, in some cases we first performed Hoxb4 immunostaining (Fig. 6B), followed by β-galactosidase histochemistry on the same transgenic embryos (Fig. 6E) and found identical patterns.

We characterized the nature of the transgene RA response in two ways. Firstly, we exposed 9.5 dpc transgenic embryos to RA in culture and then assayed for reporter staining, observing a rapid three rhombomere anterior shift in the boundary of expression (Fig. 6H). Similarly, when embryos exposed to RA at 9.5 dpc in utero were harvested and stained for reporter activity 4-6 hours later we again observed the three rhombomere anterior shift in expression (Fig. 6I). Thus, the lack of response to early RA treatments (7.5-7.75 dpc) when both the transgene and endogenous gene are expressed in the neural ectoderm does not imply the Hoxb4 gene itself is slow in responding or requires a higher dose of RA. Instead once the gene is competent to respond in later stages, the response to RA can be observed both in vitro and in utero within 4-6 hours of RA administration, suggesting that this is a rapid and direct result of RA exposure. Together these results demonstrate a conservation in the 3' location of neural-specific enhancers from three group 4 paralogs, with a close association of regulatory elements capable of mediating responses to RA that closely resemble those seen with the respective endogenous genes.

**DISCUSSION**

In this study through transgenic analysis of the human HOXD4 gene, we have identified control elements in the 3' flanking region that direct expression of a reporter gene in a pattern similar to the endogenous Hoxd4 mouse gene with respect to anterior boundaries. We found two different enhancers that independently direct mesodermal expression to the proper somite 5/6 boundary and a third regulatory element, conserved with the mouse gene, that mediates expression up to an r6/7 limit in the hindbrain. The neural enhancer involves the interaction of multiple components for r6/7 expression and also responds to ectopic doses of RA in a manner that mimics the RA response of the endogenous Hoxd4 gene. In addition, our analysis of the Hoxa4 gene identified mesodermal and neural RA-responsive regulatory elements in the 3' flanking region. These findings, and previous work on Hoxb4 and Hoxa4, have enabled us to examine the degree to which Hox group 4 regulatory elements have been conserved during duplication and divergence of the vertebrate clusters. The data show that the global organisation of cis-regulatory regions between Hoxd4, Hoxb4 and Hoxa4 is remarkably conserved, in support of the idea that they reflect the original positioning of regulatory regions in the ancestral vertebrate cluster (summarized in Fig. 7). Superimposed upon this organisation, in some cases the activity of individual elements appears to have diverged, which may account for some of the minor differences in anterior boundaries and timing of expression. In other cases activities have either been gained, lost or rearranged.

**Neural regulation of Hoxa4, Hoxb4 and Hoxd4**

Enhancers capable of setting the r6/7 anterior limit of neural expression characteristic of both the endogenous Hoxb4 and Hoxd4 genes lie in the 3' flanking sequences (Fig. 7). The timing, spatial restriction and RA response mediated by these two enhancers is very similar, suggesting that both genes may use common mechanisms to govern the pattern of neural expression. Although we used sequence comparisons to search for conservation between the Hoxb4 and Hoxd4 r6/7 enhancers, we detected no obvious block of similarity shared between them. Hence if the similar functional activity of these enhancers is derived from common ancestral regulatory elements located in this position, it would appear these activities are mediated by small dispersed motifs.

Although Hoxa4 regulation differs from that of Hoxb4 and Hoxd4, there are still important similarities. One difference is that the anterior limit of endogenous Hoxa4 expression only reaches the r6/7 junction at later time points (after 11.5 dpc) (Gaunt et al., 1989; Hunt et al., 1991) while, in earlier stages, the anterior boundary lies caudal to r6/7 (Fig. 4G,H). The 3' enhancer from Hoxa4 identified here, directed neural expression with an anterior boundary in the caudal hindbrain/spinal cord region, but the boundary never mapped to the r6/7 junction at any stage (Fig. 4A-F). This pattern closely resembles only the early phase of endogenous Hoxa4 expression. The activity of the Hoxa4 enhancer is very similar to that of the HOXD4 enhancer containing only the neural-specific component (compare Fig. 3F and Fig. 4F). Our transgenic analysis revealed that the human HOXD4 neural enhancer could be divided into two separate components; one that mediated neural specificity and a second required for setting the proper anterior boundary at r6/7 (Fig. 3E,F). This suggests that the Hoxa4 gene may differ from its paralogs in having lost the 3' component(s) required to set the r6/7 anterior limit at 10.5 dpc. On the basis of these data, we conclude that the three group 4 paralogs, Hoxa4, Hoxd4 and Hoxb4, have maintained the position of a 3' neural enhancer (see Fig. 7), presumably derived from a common ancestral regulatory element, during duplication and divergence of Hox complexes in vertebrate evolution.

**Regulation of the 3' neural enhancers by retinoids**

One of hallmarks of the Hox gene complexes is the colinear nature of their response to retinoids (Conlon and Rossant, 1992; Kessel and Gruss, 1991; Papalopulu et al., 1991; Simeone et al., 1990, 1991). However, this response is poorly
understood and it is not clear whether the RA-induced changes in expression are direct or indirect and whether they involve graded sensitivities to RA concentrations. In support of a direct role for RA and its nuclear receptors, RAREs have been found near some genes and several are implicated in the normal control of Hox expression (Dupé et al., 1997; Frasch et al., 1995; Langston and Gudas, 1992; Marshall et al., 1994; Moroni et al., 1993; Morrison et al., 1996; Popperl and Featherstone, 1993; Studer et al., 1994). The RA response of the most 3' genes has previously been examined in the greatest detail, but our analysis of three group 4 Hox genes has revealed that they respond in a different manner. RA treatment at times (7.25-7.75 dpc) known to affect Hoxb1 expression (Conlon and Rossant, 1992; Marshall et al., 1992) had no effect upon the endogenous expression pattern of Hoxa4, Hoxb4 or Hoxd4 even though these genes are being expressed in the early neural ectoderm and establishing their respective anterior boundaries closely following Hoxb1. Conversely, RA treatment at later times (8.5-9.5 dpc) rapidly induced anterior shifts in expression of the group 4 genes (Figs 5, 6), when Hoxb1 is no longer responsive even though it is still expressed in the hindbrain.
hours of exposure. To RA at 9.5 dpc, illustrating that the response occurs within 4-6 rapid RA response is seen in a 9.75 dpc transgenic embryo exposed rhombomere anterior shift of expression is observed. (I) The same embryo exposed transiently to RA in embryo culture, again a three respond to RA at this time. (H) An ~10.5 dpc Hoxb4 (F). (G) A 9.5 dpc transgenic embryo exposed to RA at 7.5 dpc. The Hoxb4 transgene, like the endogenous gene (not shown), fails to respond to RA at this time. (H) An ~10.5 dpc Hoxb4 transgenic embryo exposed transiently to RA in embryo culture, again a three rhombomere anterior shift of expression is observed. (I) The same rapid RA response is seen in a 9.75 dpc transgenic embryo exposed to RA at 9.5 dpc, illustrating that the response occurs within 4-6 hours of exposure.

Both Hoxb4 and Hoxd4 displayed identical rapid rostral shifts in expression of one rhombomere at 8.5 dpc. Similarly, at 9.5 dpc, they had an anterior shift of three rhombomeres (r6-r4) in the limits of expression, but the induction is dynamic and followed by a specific down regulation in r5. This suggests that there are differences in the molecular environment of r5 which are incompatible with sustained expression of the Hoxb4 and Hoxd4 genes, and might reflect a basic difference between odd- and even-numbered rhombomeres. The ability of the 3’ enhancers to mediate a response to RA and r6/7 neural expression that resembles the endogenous Hoxb4 and Hoxd4 genes, provides further support for the idea that both genes use common mechanisms in directing neural expression in the hindbrain (Gould et al., 1997). With respect to the RA response, it remains unclear if this is directly mediated by retinoid receptors, because there are no typical consensus RARE direct repeats (DR2 or DR5) in the enhancers, but there are several dispersed half sites.

The temporal window of the RA response of Hoxa4 was similar to that of Hoxb4 and Hoxd4, but the nature of the response was very different. While normal Hoxa4 expression does not extend up to the r6/7 junction prior to 10.5 dpc, RA treatment prematurely resulted in r6/7 expression but there was no induction in more anterior rhombomeres. The neural enhancer that we identified 3’ to the Hoxa4 also responds to RA in a similar manner to the endogenous gene, but the response was restricted only to the ventral regions of the neural tube. Hence, the Hoxa4 3’ enhancer is only capable of mediating part of the overall endogenous neural expression and RA response.

Thus, three members of the Hox group 4 paralogs and their 3’ neural enhancers have a conserved temporal window of sensitivity to exogenous retinoids that opens around 8.5-9.0 dpc and may be related to their positions in the Hox complexes. This is also similar to the RA response that we observed with a 5’ HOXD4 neural enhancer containing an RARE (Morrison et al., 1996). Once these genes are competent, their response to retinoids is rapid (within 4-6 hours in all cases). Indeed it would appear that, once a group 4 gene is able to respond to RA, it is as sensitive as the most 3’ genes, Hoxb1 or Hoxa1. These results suggest that the colinear response of Hox genes to retinoids is not a simple function of variable thresholds or graded concentrations along the A-P axis as previously suggested (Dekker et al., 1992; Gaunt and Strachan, 1994, 1996; Simeone et al., 1991). Other cofactors or inhibitors appear to be involved in limiting the ability of the genes to respond directly or indirectly to RA. The molecular basis for these temporal windows of RA sensitivity is not known but it has been suggested that vertebrate members of the trithorax and polycomb groups might be involved in controlling the nature of the RA response and patterns of Hox genes (Coré et al., 1997).

Mesodermal regulation of Hoxa4, Hoxb4 and Hoxd4

Despite their coincident expression up to the r6/7 junction in the hindbrain, the three group 4 paralogs, Hoxa4, Hoxb4 and Hoxd4, have anterior boundaries of expression in somitic mesoderm that are offset from each other by a single somite. Previous analysis of Hoxb4 indicated that the intron contained an enhancer that directed expression up to s6/7 (Whiting et al., 1991). Here we have shown that the 3’ flanking regions of the Hoxd4 and Hoxa4 genes contain mesodermal enhancers able to direct somitic expression which correspond to the endogenous genes (Fig. 7). In the case of Hoxd4, there are two separate mesodermal enhancers flanking the neural element, each capable of mediating expression up to the proper s5/6 junction (Fig. 2F-M). For Hoxa4, we have mapped a single
mesodermal enhancer 5' to the neural element which directs expression up to the s7/8 junction. Regions 3' to the neural element were not examined, so by analogy to HOXD4 it remains possible that there are additional mesodermal control regions. The similar organisation between 3' mesodermal control regions of HOXD4 and Hoxa4 led us to re-examine Hoxb4, and we also detected posterior somitic expression in a line containing the 3' region A neural enhancer and have mapped a separate mesodermal control region flanking the minimal neural enhancer element CR3 (Fig. 7; unpublished observations). Therefore, these three group 4 paralogs have similarly positioned 3' mesodermal enhancers, but each with a slightly different anterior boundary in somites. Sequence comparisons between the different mesodermal enhancers have not identified any extended blocks of sequence conservation but, if they are evolutionarily related, they may depend on a number of small dispersed motifs for activity.

Extended paralog comparisons

Given the apparent similarity in organisation of 3' regulatory regions described here to what extent are additional control elements in the Hoxa4, Hoxb4 and HOXD4 loci also conserved? Transgenic studies have mapped mesodermal and neural components 5' of Hoxa4 (Behringer et al., 1993) that are involved in regulating its expression. A 5' neural region is required for late phase expression of Hoxa4, which reaches the r6/7 junction, whereas the 3' region identified here mediates the earlier pattern. In contrast, the 5' and 3' mesodermal regions appear to mediate similar somite expression up to s7/8. Therefore, there is considerable overlap in the regulatory activities of the Hoxa4 5' and 3' enhancers (Fig. 7).

In the case of HOXD4, we previously characterised a 5' enhancer containing an RARE that directed neural restricted expression and a response to RA (Morrison et al., 1996). Recently, in a parallel study of the mouse Hoxd4 5' flanking regions, neural activity similar to that seen with the human HOXD4 gene, and a mesodermal enhancer activity associated with the 5' RARE was identified (Zhang et al., 1997). In further support of a general positioning of 5' control elements in the group 4 genes, we have begun to analyse the 5' flanking regions of the Hoxb4 gene and also found neural and mesodermal enhancer activities that direct expression with anterior boundaries caudal to the endogenous gene (Fig. 7; S. Nonchev, J. Sharpe and R. K., unpublished). In addition to the 5' and 3' control regions, the intron of Hoxb4 contains an enhancer (region C) with both neural and mesodermal activities (Whiting et al., 1991). This enhancer is capable of mediating mesodermal expression directed up to the s6/7 junction observed with the endogenous gene, neural domains with an anterior limit in the caudal hindbrain/spinal cord region. The presence of a highly conserved region (CR1) within the introns of Hoxb4, Hoxd4 and Hoxa4 in broad range of vertebrate species (Aparicio et al., 1995; Cianetti et al., 1990; Morrison et al., 1995) suggested that there might also be conserved regulatory activities in the introns of all the group 4 paralogs (Fig. 7). However, our analysis here on the HOXD4 intron and CR1, and previous analysis on Hoxa4 (Behringer et al., 1993), indicates that in contrast to Hoxb4 neither of these introns have independent enhancer activities. Therefore, if CR1 has a general role in the group 4 paralogs, it may be through interaction/cooperation with additional 5' or 3' regulatory regions.

Fig. 7 summarises the relative positioning of regulatory activities in three group 4 paralogs. On the basis of the similar 5' and 3' distribution of neural and mesodermal enhancers and retinoid-responsive regions, it appears that there has been considerable conservation in the position of regulatory components of the group 4 paralogs during the duplication and divergence of the ancestral vertebrate Hox complex. Mesodermal regulation appears to be complex and variable between paralogs. The identification of multiple mesodermal-specific enhancers, each with the ability to regulate or impose the same or slightly different anterior limits, provides a future basis for examining the mechanisms that set anterior somite boundaries of Hox genes. With respect to neural control, the timing of segmental expression and organisation of neural regulatory components of Hoxb4 and HOXD4 are very similar. Furthermore, as both the endogenous Hoxb4 and HOXD4 genes are required for boundary maintenance of the Hoxb4 r6/7 neural enhancer (Gould et al., 1997), this para-regulation (i.e. cross-regulation between paralogs) indicates the segmental regulatory mechanisms of these two genes are also linked. Since the expression of Hoxd4 overlaps with Hoxb4 and Hoxd4 in regions posterior to the r6/7 boundary, cross talk involving all three of these paralogs may be critical for maintaining neural expression.

We thank Drs Dado Boncinielli, Debra Wolgemuth and Denis Duboule for providing Hox clones; Dr Stefan Nonchev performing transgenic analysis of CR1 mutants; Drs Feng Zhang, Heike Popperl and colleagues for communicating HOXD4 results prior to publication;


(Accepted 20 June 1997)