The EMBO Journal vol.12 no.9 pp.3539-3550, 1993

Structure and activity of regulatory elements involved in the activation of the *Hoxd-11* gene during late gastrulation

Matthieu Gérard[1], Denis Duboule[1,2] and József Zákány[1]

European Molecular Biology Laboratory, Meyerhofstrasse 1, Postfach 10.2209, 6900 Heidelberg 1, Germany
[1]Present address: Department of Zoology, University of Geneva, Sciences III, Quai Ernest Ansermet 30, 1204 Geneva, Switzerland
[2]Correspondence

Communicated by D.Duboule

We have used reporter gene constructs to study the *cis* regulation of the *Hoxd-11* gene (previously *Hox*-4.6) in transgenic mice. We identified a 5 kb regulatory unit, which was able to reproduce important aspects of the initial activation of the gene along the major body axis. The comparison of the nucleotide sequence of this DNA fragment with the corresponding avian genomic region revealed the presence of seven highly homologous stretches of DNA outside the protein coding regions. In particular, the 3′ flanking region contained two such domains that are required to mediate the embryonic activation. A chimeric construct containing the two short homologous regions from the chicken gene could replace the complete murine fragment thus demonstrating that the conserved domains carry the main regulatory elements involved in this activation. The first half of this bipartite regulatory region has enhancer activity when tested with a heterologous promoter, while the second half is required to restrict the enhancer activity to the proper expression domain. These results suggest that stage- and tissue-specific cooperation between regulatory elements is required to control properly the activity of the *Hoxd-11* promoter.

Key words: colinearity/development/gene regulation *in vivo*/homeobox/*HoxD* complex

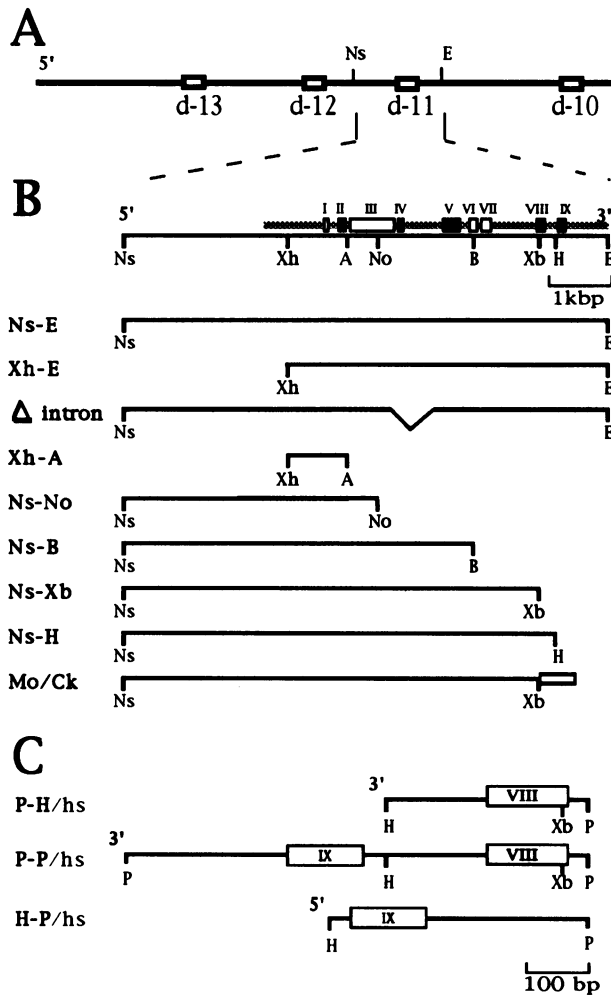
Introduction

The mammalian Class 1 homeobox (*Hox*) genes are members of a family of 38 related genes which are clustered in four complexes, *HoxA*, *B*, *C* and *D*, located on different chromosomes (Scott, 1992). The regulation of *Hox* gene expression during development strictly reflects the location of individual genes in the complexes as their 3′ to 5′ order determines their sequence of expression along the anterior-posterior main body axis (structural colinearity: Gaunt *et al.*, 1988; Duboule and Dollé, 1989; Graham *et al.*, 1989). A possible reason for this anterior to posterior sequence of *Hox* expression domains is the differential temporal activation of these genes during morphogenesis which progresses towards more posterior body regions (temporal colinearity: Izpisúa-Belmonte *et al.*, 1991b). The same features in the functional organization of *Hox* genes have also been observed during pattern formation along other body axes such as the limb (Dollé *et al.*, 1989; Izpisúa-Belmonte *et al.*, 1991a; Yokouchi *et al.*, 1991) and genital axes (Dollé *et al.*, 1991a), suggesting that the *Hox* gene family may play a similar role in all axial structures (Duboule, 1992). Recent genetic manipulations involving loss of function (Chisaka and Capecchi, 1991; Lufkin *et al.*, 1991; Chisaka *et al.*, 1992; LeMouellic *et al.*, 1992) and gain of function (Kessel *et al.*, 1990; Jegalian and De Robertis, 1992; Lufkin *et al.*, 1992; Morgan *et al.*, 1992; Pollock *et al.*, 1992) experiments indicate that *Hox* gene products are important for normal development as perturbation of their expression leads to severe morphological defects. Hence normal morphogenesis is critically dependent upon the proper establishment of *Hox* gene expression domains and it thus becomes important to understand the molecular mechanisms that coordinate their sequential activation in space and time.

Hox gene transcripts usually accumulate from a specific body level towards the tail, from about the time at which this anterior-most region forms. Some important aspects of the control of *Hox* gene activation could be reproduced when genomic fragments were placed outside of the *Hox* complexes, in transgenic mice. For example, transgene expression patterns obtained with the *Hoxa-7* (Püschel *et al.*, 1991) and *Hoxb-4* (Whiting *et al.*, 1991) large genomic fragments clearly resemble that of the endogenous genes and thus validate the use of this approach to search for *cis*-acting regulatory elements. However, precisely defined regulatory elements that could control time-restricted initial activation have not yet been characterized. As the specific sequence of gene activation might be, in part, dependent upon interactions *in cis* between different control elements, the identification and analysis of such elements as well as their interactions is an essential step in understanding the functional organization of *HOX* complexes.

We used the posterior part of the *HoxD* complex as a model system and were able to show that a DNA fragment, including the entire *Hoxd-9* gene, was sufficient to generate a rather faithful, posteriorly restricted, expression pattern in transgenic mouse embryos (Renucci *et al.*, 1992). *Hoxd-9* is one of the five neighboring genes which are located at the 5′ end of the complex and which are related to the *Drosophila Abd-B* homeotic gene. Of these five consecutive genes *Hoxd-9* is the first to be activated in 8 day old mouse embryos. It is followed by the four other genes, in a temporal order that corresponds to their position so that the genes are progressively activated, from *Hoxd-9* to *Hoxd-13*, during the formation of the most posterior parts of the body (Izpisúa-Belmonte *et al.*, 1991b). In order to compare the mechanisms of activation among these highly coordinated genes, we set out to analyze the *cis*-acting elements controlling the expression of the *Hoxd-11* gene.

In this report, we describe a complex regulatory unit in the *Hoxd-11* region. This region was able to mimic, in both space and time, the restricted activation of the endogenous



D Incidence of expression patterns

	Late patterns				Early patterns			Total
	Trunk s45	Limb	Abdomen	Genital eminence	Genital bud	Limb bud	Trunk s27	β-gal/lacZ
Ns-E	4/7	8/10	9/10	0/10	7/7	8/9	11/12	12/16
Xh-E	nt.	0/6	0/6	0/6	3/6	2/6	6/7	7/11
Δ intron	1/1	5/6[a]	5/6	0/6	6/6	6/6	6/6	6/8
Xh-A	nt.	nt.	nt.	nt.	0/4	0/4	0/4	4/7
Ns-No	2/4	4/4	3/4	3/4	0/4	0/4	0/4	4/6
Ns-B	0/2	4/4	4/4	0/4	0/4	0/4	0/4	4/6
Ns-Xb	0/3	0/4	0/4	0/4	0/4	0/4	0/4	4/3[b]
Ns-H	nt.	3/5	3/5	0/5	4/5	5/9	7/10[a]	10/11
Mo/Ck	1/1	3/5	4/5	0/5	5/5	4/5	5/5	5/8
P-H/hs	nt.	nt.	nt.	nt.	nt.	3/3[a]	6/7[a]	7/7
P-P/hs	nt.	nt.	nt.	nt.	nt.	0/2	2/2	2/4
H-P/hs	nt.	nt.	nt.	nt.	nt.	0/2	0/2	2/5

Fig. 1. Characterization of regulatory regions of the *Hoxd-11* gene in transgenic mouse embryos. Summary of the constructs and sites of expression. (**A**) Scheme of the posterior part of the murine *HoxD* complex with the distances between the homeoboxes of the different genes (rectangles) and the approximate positions of the *Nsi*I (Ns) and *Eco*RI (E) restriction sites which delimit the genomic fragment analyzed in this study. (**B**) A partial restriction map of the *Hoxd-11* gene is shown below the horizontal line which represents the *Nsi*I−*Eco*RI genomic fragment. The rectangles numbered I−IX above show the position of the regions of interspecies conservation (open or black boxes represent a minimum of 70% or 85% nucleotide identity, respectively). The dotted line across the boxes indicates the sequenced region. A, *Apa*LI; B, *Bam*HI; E, *Eco*RI; N, *Nsi*I; No, *Not*I; X, *Xho*I. The different variants of the transgene constructs are shown underneath with their designations on the left. The open box attached at the end to the Ns-Xb fragment in the Mo/Ck construct represents the *Xba*I−*Apa*I 550 bp chicken *Hoxd-11* fragment. (**C**) Schematic representation of the *Hoxd-11* genomic sequences tested in the enhancer assay. The 330 bp *Pst*I−*Hin*dIII, 730 bp *Pst*I−*Pst*I and 400 bp *Hin*dIII−*Pst*I fragments were placed in front of the mouse *hsp68* promoter. The direction of transcription is from left to right, into the *lacZ* gene. The 5′−3′ orientation of the fragments was reversed in the P-H/hs and P-P/hs constructs with respect to the *hsp68* promoter. (**D**) Results of transgene expression analysis. The incidences of specific expression patterns are shown aligned with the respective constructs in B and C. The 'early' domains were established by embryonic day 10 in synchrony with the appearance of the respective organ primordia. Expression in the 'late' domains was detected subsequently (see Results for details). The last column lists the total number of β-gal positive individuals, including those with expression patterns not related to the reproducible *Hoxd-11* specific patterns, over the total number of *lacZ* positive independent integrants, as determined by Southern blotting analysis of genomic DNA samples. The box highlights the lack of early activation in the absence of regions VIII and IX. [a]The expression patterns observed were more extensive than that of the original Ns-E construct, s25 and s16 in the cases of Ns-H and P-H/hs, respectively. [b]The genomic DNA isolated from the extra-embryonic membrane of an otherwise clearly β-gal positive embryo was not *lacZ* positive in Southern analysis. nt, not tested.

gene along the main body axis, together with most of its other activation characteristics. Although the transgene expression pattern in developing limbs was not the same as that of the endogenous *Hoxd-11* gene, selective activation in the posterior mesoderm of both forelimbs and hindlimb buds was observed, in a way reminiscent of the expression domain of the endogenous gene. We compared the nucleotide sequence of the mouse locus with that of the corresponding chicken genomic region (Izpisúa-Belmonte *et al.*, 1991a). The regulatory potential of different combinations of the highly conserved sequence domains was analyzed *in vivo*, which led to the isolation of a short bipartite regulatory region required for the embryonic activation of the *Hoxd-11* transgenes. The analysis of this region suggests that it is involved in the control of the initial activation of the *Hoxd-11* gene in the posterior body region as well as in limbs and the genital bud.

Results

Developmental expression of a Hoxd-11−lacZ fusion gene

We first analyzed the expression pattern of a large construct that contained approximately 3 and 2 kb of sequence upstream and downstream of the *Hoxd-11* transcription unit, respectively (Ns-E; Figure 1A and B and Figure 2). The

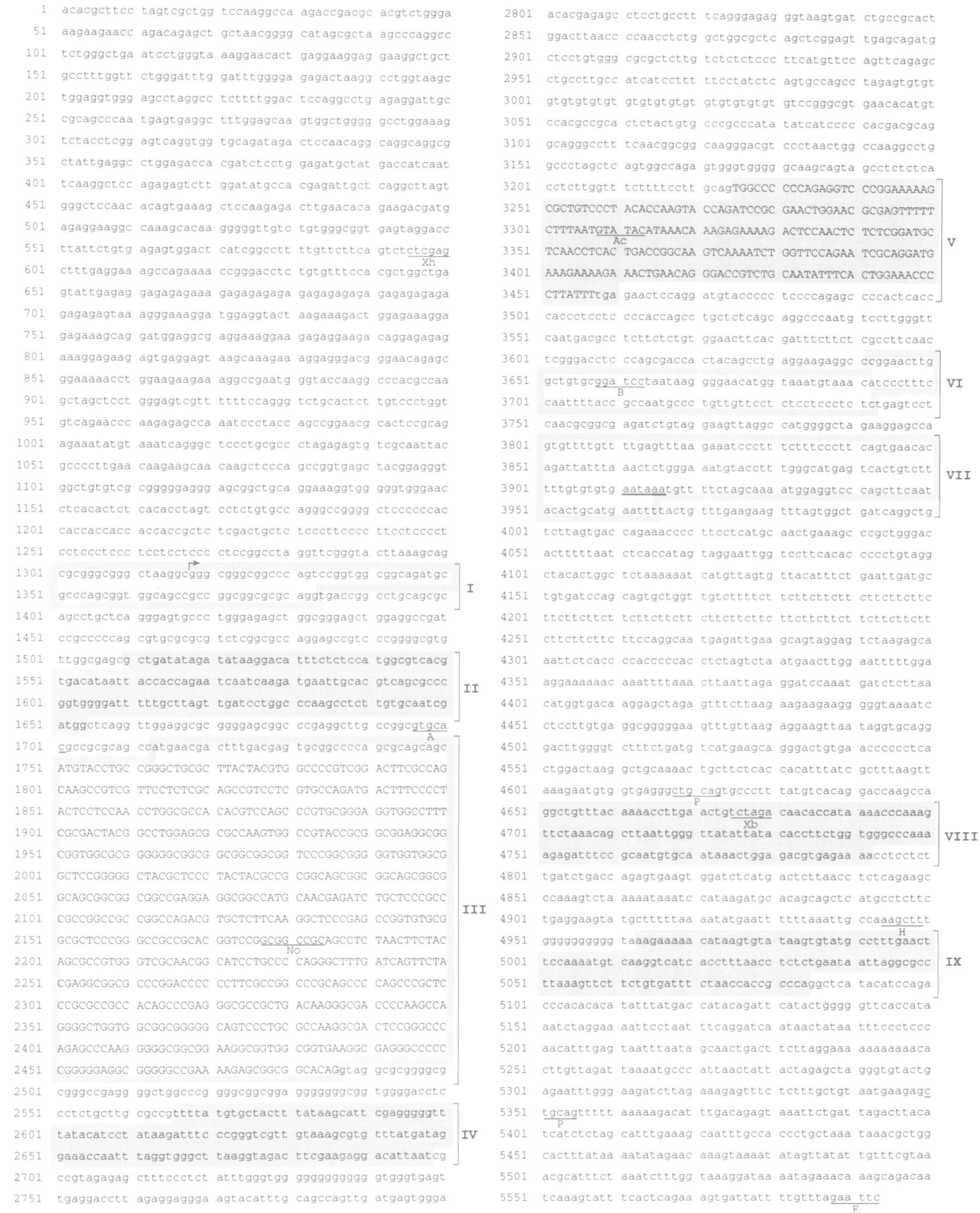


Fig. 2. Nucleotide sequence of the mouse *Hoxd-11* locus. The nine conserved regions are numbered from I to IX. Open boxes (I, III, VI and VII) indicate 70% identity between mouse and chicken whereas shaded boxes (II, IV, V, VIII and IX) indicate a minimum of 85% identity. The arrow shows a major transcription start site. The polyadenylation site as well as the restriction sites used to generate the constructs shown in Figure 1 are underlined. A, *Apa*LI; Ac, *Acc*I; B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; No, *Not*I; P, *Pst*I; Xh, *Xho*I; Xb, *Xba*I. Coding regions are shown in capital letters (for the protein sequence, see Izpisúa-Belmonte *et al.*, 1991b). This sequence is available from EMBL/GenBank/DDBJ under accession number X71922.

production of transgenic families allowed us to analyze in detail the establishment of the transgene expression pattern during development. The time of activation along the main body axis was on embryonic day 9 (E9). At this stage, the four occipital somites were no longer discernible. In the following description, we have therefore numbered only

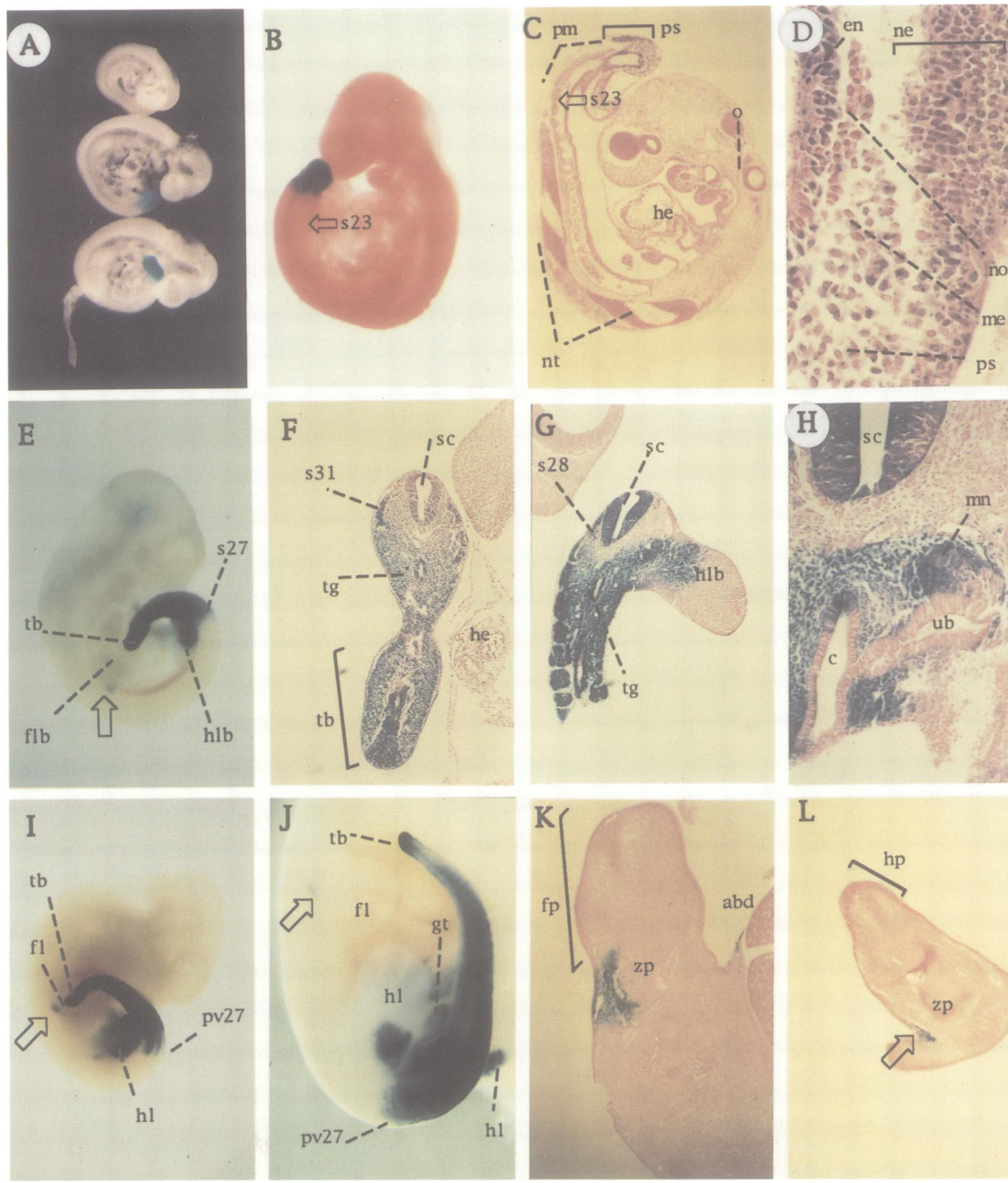


Fig. 3. Early activation and development of the *Hoxd-11* transgene expression pattern mediated by the Ns-E construct. (**A**) Initial activation at the posterior ends of three E9 embryos (Theiler stage 14–15) sired by a homozygous male of family #397. Activity is not seen in the top embryo (~18 somites) but appears in the one in the middle panel (~23 somites) and becomes broader in the bottom panel (~25 somites). (**B**) A 23 somite stage embryo after eosin staining. The arrow points to the level of the last condensed somite. (**C**) Histological section taken from an equivalent stage Xh-E embryo which shows the presence of histochemically detectable β-gal in and around the regressing primitive streak. (**D**) Enlargement of a parallel section of the same embryo which shows the posterior end. β-gal activity is detected in all cell types. At this stage, the Ns-E or Xh-E fragment-mediated expression patterns are indistinguishable. (**E**) Expression pattern of an Ns-E transgene at E10 (Theiler stage 16–17). Note the expression in the hindlimb bud, the posterior, proximal forelimb bud and the tail. Weak expression in the head mesoderm is specific for this family and was not observed in the other families or embryos obtained with the Ns-E construct. (**F**) Histological section through the tail and tailbud of an E10 embryo shows that derivatives of all three germ layers are β-gal positive. (**G**) Histological section through the posterior hindlimb bud of an Ns-E embryo at an equivalent stage. (**H**) A parallel section of the same embryo shows the strong signal in the metanephric anlage, and absence of β-gal activity in the ureteric bud. (**I**) Pattern of transgene expression of the Ns-E construct in an E11 (Theiler stage 19) embryo. (**J**) Whole mount preparation of an E12.5 (Theiler stage 21) Ns-E embryo showing the posterior expression domains in the trunk, forelimbs, hindlimb, tail, genital bud and abdominal wall. (**K** and **L**) Histological sections through the forelimb and hindlimb, respectively, of an Ns-E late E12 embryo (Theiler stage 20–21). The open arrows in panels E, I and L point to the β-gal positive posterior-proximal mesoderm cells in the forelimbs. c, cloaca; en, endoderm; fl, forelimb; flb, forelimb bud; fp, footplate; gt, genital tubercle; he, heart; hl, hindlimb; hlb, hindlimb bud; hp, handplate; me, mesoderm; mn, metanephros; no, notochord; ne, neuroectoderm; nt, neural tube; o, otic vesicle; pm, pre-somitic mesoderm; ps, primitive streak; s23, s31, somite number 23 and 31, respectively; sc, spinal cord; tb, tailbud; te, telencephalon; tg, tailgut; ub, ureteric bud; zp, zeugopod skeletal blastema.

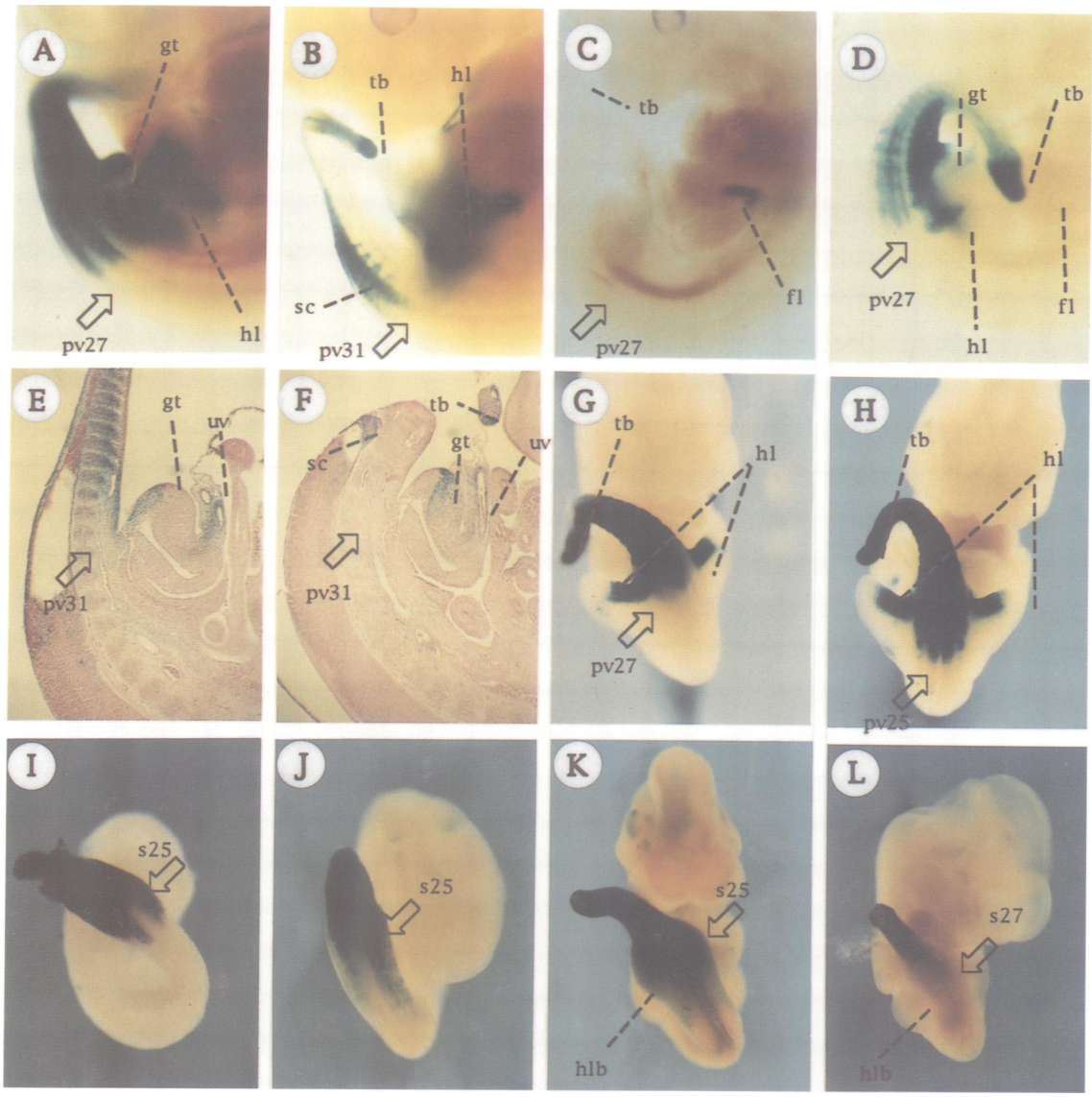
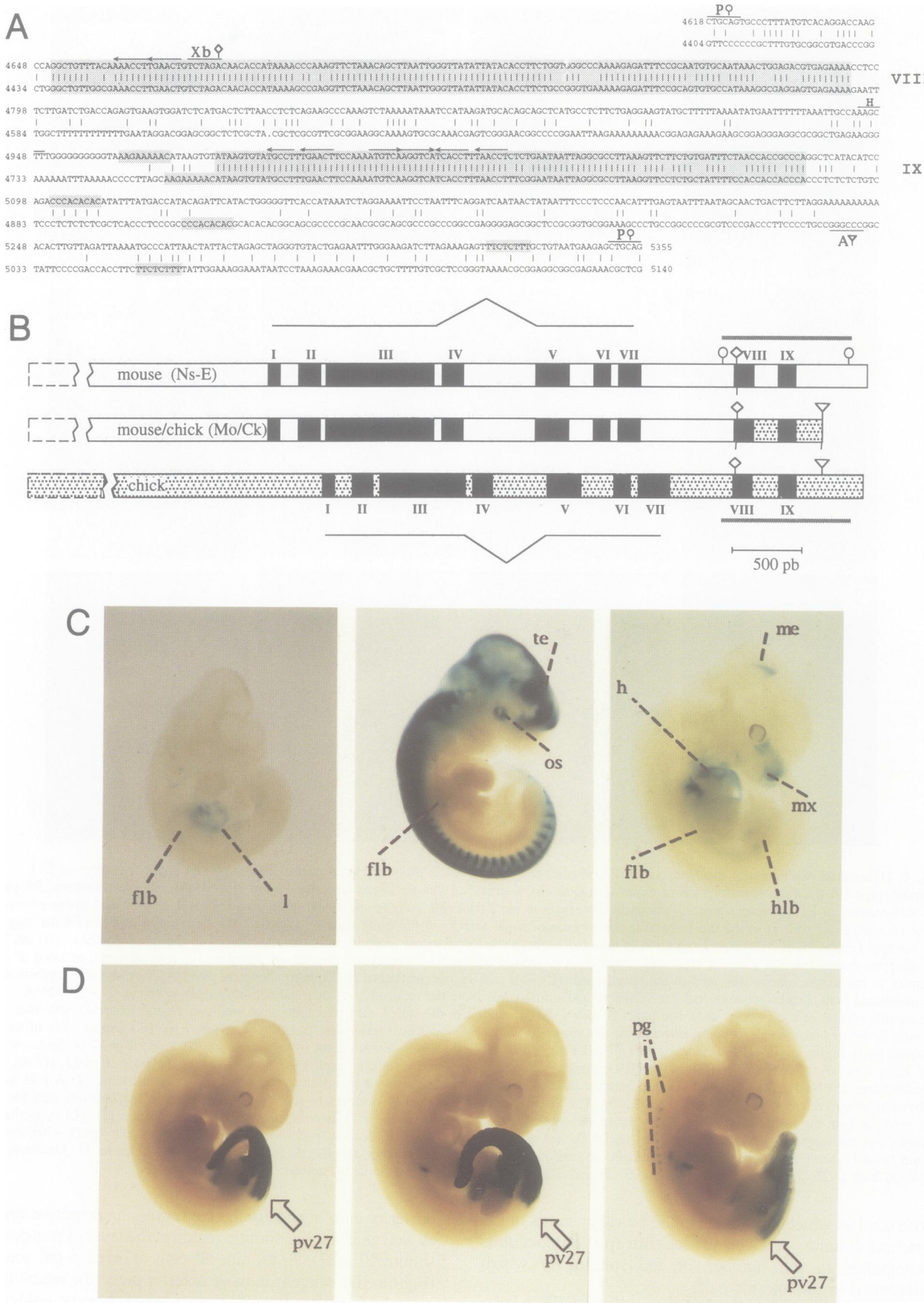


Fig. 4. Differential expression of the *Hoxd-11* Ns-E construct and derivatives in the posterior region of the body. (**A**) A Ns-E embryo at E11 (Theiler stage 19) shows the combination of expression sites characteristic of this transgene (see also Figure 3I). (**B**) An embryo transgenic for the Ns-No construct at E11 (Theiler stage 19) showing expression in the dorsal root ganglia and spinal cord from the level of pv31 and in the extreme posterior somites as well as in the tail bud. Note the absence of β-gal activity in the trunk anterior to pv31. (**C**) An embryo at E11 (Theiler stage 19) positive for the Ns-B construct. Expression is detected in the posterior-proximal forelimb but is much reduced in the other domains. (**D**) An Xh-E transgenic embryo at E11 (Theiler stage 18). Expression is seen in all the Ns-E embryonic expression domains, but appears much reduced in intensity in the somitic derivatives, and in the spinal cord. (**E**) This mid-sagittal section of the embryo shown in panel A illustrates the expression in the abdominal wall, genital tubercle, tail pre-vertebrae and spinal cord. Note the reduced level of β-gal activity in pre-vertebrae 27−30. (**F**) A comparable section of the embryo shown in panel B shows the expression in the spinal cord, the abdominal wall and genital tubercle. Note that staining in the genital tubercle is complementary to that obtained with Ns-E (or Xh-E) transgenic embryos (e.g. panel E). (**G**) Dorsal view of the posterior body region of an Ns-E embryo at E11 (Theiler stage 19). Note the sharp anterior limit of expression at the level of pv27 in the trunk. (**H**) The same view of an Ns-H E11 (Theiler stage 19) embryo. Note the sharp anterior limit of expression in the trunk at the level of pv25. (**I**) An E9 (Theiler stage 15) Ns-H embryo with the sharp anterior limit of expression at the level of prospective somite 25 in the neural tube. (**J**) A P-H/hs embryo at an equivalent stage (E9; Theiler stage 15). The arrow points to the level of prospective s25. β-gal is detected in tissues more anterior than this level, in both the neural tube and the mesoderm. (**K**) An embryo transgenic for the P-H/hs construct at E10 (Theiler stage 17). (**L**) A similar embryo, but transgenic for the P-P/hs construct (E10; Theiler stage 17−18). Note the reduced level of expression in the anterior (pv27−30) sector of the *Hoxd-11* Ns-E embryonic expression domain. ab, abdominal wall; drg, dorsal root ganglia; fl, forelimb; gt, genital tubercle; hl, hindlimb; hlb, hindlimb bud; pv25, 27 and 31, pre-vertebrae 25, 27 and 31, respectively; sc, spinal cord; tb, tailbud; uv, umbilical vessels.

those somites which participate in the morphogenesis of the vertebral column. A given prevertebra (pv) number thus corresponds to the number of the somite (s) which contributed to its anterior half.

Whole mount and histological analysis revealed that, in the 23 somite stage embryo, the anterior body regions and most of the pre-somitic mesoderm were negative (Figure 3A−C). β-gal activity was first detected in the posterior end, in and around the regressing primitive streak, as well as in the pre-somitic region at the level of prospective somite 27 (s27). In the tip, all cells showed β-gal activity (Figure 3D) whereas in more anterior parts, the neural tube, mesoderm and endoderm were positive but the epidermis was negative.

At E10, expression was detected in the somites and their derivatives, and in the neural tube and neural crest, posterior


A
Xb
P
H
VIII
IX
B
mouse (Ns-E)
mouse/chick (Mo/Ck)
chick
I II III IV V VI VII VIII IX
500 pb
C
flb
l
te
os
me
h
mx
hlb
D
pv27
pg

to s26 (Figure 3E–G). However, both the hindgut endoderm and the ureteric bud which at this stage is in intimate contact with the strongly positive metanephric anlage (Figure 3H), were negative. In the tail, all tissues were positive, with a stronger signal in the ventral parts of the tailgut, tail mesoderm and neural tube. β-gal activity was also detected in the posterior third of the hindlimb bud mesoderm, with a higher intensity dorsally (Figure 3G). A posterior-ventral expression domain was subsequently established in the forelimb, which slightly extended into the ventro-lateral trunk in older specimens (Figure 3E and I–L; open arrow).

By E11.5, the pattern extended into the genital tubercle and the umbilical as well as peri-umbilical abdominal mesoderm (Figure 3I and J). A few hours later, expression at the level of pre-vertebrae 27–30 (pv27–30) faded considerably in all cell types with the exception of the spinal cord (Figure 4A and E). At this developmental stage, the apparent anterior limit in the axial mesoderm was at pv31. This expression pattern remained stable at least until E13.5.

Definition of a regulatory unit in the Hoxd-11 locus

We tried to reproduce this rather faithful pattern of embryonic activation (observed with the Ns-E construct) by using shorter genomic fragments. A deletion analysis, involving 3′, 5′ and intronic sequences, was therefore carried out (see Figure 1 for the various constructs). A large 5′ deletion construct (Xh-E; Figure 1B) still resulted in expression patterns similar, at early stages, to that of the original Ns-E transgene. The activation of expression was reproduced at E9 (Figure 3B) and histological analysis revealed a comparable distribution of β-gal positive cells. Expression was observed in all cells posterior to s26 in the unsegmented mesoderm, neural tube, and in the endoderm (Figure 3C and D). In older embryos (Figure 4D) expression was found in the tissues posterior to pv26, in the spinal cord, pre-vertebrae, dorsal root ganglia, tail, tailbud and metanephros. However, similarly to the Ns-E construct, expression was detected neither in the ureteric bud nor in the endodermal layer of the hindgut. In four independently derived embryos, the anterior limit of expression was found at pv27. In three embryos, expression in the posterior part of the genital tubercle was also visible (Figure 4D). These embryos were clearly old enough for one to expect to see expression in the proximal-posterior forelimb domain as well as in the umbilical and abdominal mesoderm. However, expression in these structures was not detected.

These results suggested that the genomic sequences present in the Xh-E construct are sufficient for a fairly normal activation of the transgenes in the early domains of expression along the main body axis, the hindlimb bud and the genital tubercle. This combination of expression domains defined an early embryonic pattern that resembled that of the endogenous *Hoxd-11* gene. We therefore focused on a more detailed analysis of this 5 kb *Xho*I–*Eco*RI region.

Highly conserved DNA regions in vertebrate Hoxd-11 loci

The nucleotide sequence of the mouse *Xho*I–*Eco*RI fragment was compared with that of the corresponding chicken *Hoxd-11* genomic region. Figure 2 shows the nucleotide sequence of the mouse locus with the regions that showed significant sequence homology with the avian locus (see also Figure 5A). Nine regions of particularly high sequence similarity were found (I–IX) inside as well as outside the *Hoxd-11* coding region. Regions I and VII were found to delimit the *Hoxd-11* transcription unit. Thus, while region I (85 bp, 69% identity) contained a major transcription start site (localized by RNase mapping, arrow at position 1317 in Figure 2), region VII (211 bp, 72%) contained the mRNA polyadenylation signal, AATAAA (starting at position 3911 in Figure 2). The *Hoxd-11* mRNAs, identified by Northern blot analysis, were ~2 kb and 1.8 kb long for the mouse (Izpisúa-Belmonte *et al.*, 1991a) and chick (Rogina *et al.*, 1992) respectively. This agrees well with the distances between regions I and VII, when the lengths of the introns (870 bp and 830 bp for mouse and chicken, respectively) were subtracted.

Five additional domains of high interspecies sequence similarity were uncovered in the *Hoxd-11* transcription unit. Region II (161 bp, 89%) was found in the 5′-untranslated sequence, region III (71%) matched with the protein coding part of the first exon, whereas region IV (141 bp, 84%) was

Fig. 5. Identification of an early embryonic activating region located at the 3′ end of the *Hoxd-11* gene and conserved between birds and mammals. (**A**) Alignment of the mouse *Hoxd-11* 3′ nucleotide sequence (nucleotides 4622–5354 in Figure 2) with the corresponding chicken sequence. Shaded boxes highlight conserved regions VIII and IX. Scattered shaded sequences show motifs which are conserved but whose relative positions have changed. The arrows indicate motifs related to the consensus DNA binding sites for some nuclear receptor molecules. The restriction sites used for transgenic analysis are indicated. A, *Apa*I; H, *Hin*dIII; P, *Pst*I; Xb, *Xba*I. (**B**) Schematic representation of the Mo/Ck chimeric construct. The upper line shows the mouse genomic region with the positions of the exons shown above. The black boxes represent the conserved regions I–IX. The bottom line represents the same illustration but for the chicken counterpart (stippled). The dotted lines above and below the illustrations of the mouse and chicken genes show the position of the DNA sequences shown in panel A. The chimeric construct is shown in between (Mo/Ck transgene). In this construct, the mouse terminal part was exchanged against the corresponding chicken one. The symbols in this region represent the position of the restriction sites, as indicated in panel A. (**C**) Three transgenic embryos are shown that express the Ns-Xb transgene (see Figure 1). The left panel is an E10 embryo (Theiler stage 16–17) that shows staining in the anterior and posterior margins of the branchial arches, a double band of cells in the myelencephalon, in the lung mesoderm, in the forelimb and in the thoracic intermediate plate mesoderm. The staining intensity is very weak. The middle panel shows a E11 transgenic embryo (Theiler stage 18–19). Expression is restricted to the dorsal half of the central nervous system, at all levels, including the optic stalk and the dorsal axial mesoderm all along the central nervous system. The embryo on the right is at E11.5 (Theiler stage 19–20) and shows another abnormal set of expression domains in the maxillary processes, mesencephalon, anterior and proximal parts in both the forelimbs an the hindlimbs, heart and neural tube floorplate in the tail region, and at the anterior and posterior margins of the genital tubercle. (**D**) Three transgenic embryos that express the chimeric Mo/Ck transgene. The embryos are at E10.5 (Theiler stage 17–18; left panel), E11 (Theiler stage 18–19; middle panel) and E12 (Theiler stage 20; right panel). All three embryos have a distribution of β-gal that corresponds to that observed in Ns-E transgenic embryos of a comparable age. A slight ectopic expression is seen in the chain of cervical sympathetic ganglia in the embryo on the right. The short chicken DNA fragment is able to rescue the highly irregular expression pattern produced by the Ns-Xb transgene. The open arrows point to the anterior limit of transgene activation at the level of pre-vertebra 27. ab, abdominal wall; flb, forelimb bud; h, heart; hlb, hindlimb bud; l, lung; me, mesencephalon; mx, maxillary process; os, optic stalk; pg, pre-vertebral ganglia; te, telencephalon.

found in the intervening sequences. Region V (86%) corresponded to the homeobox region, and region VI (109 bp, 69%) was located in 3′-untranslated sequences. Regions VIII (140 bp, 87%) and IX (120 bp, 89%) were mapped 3′ to the termination of the *Hoxd-11* transcription unit, in both species (see also Figure 5A). In order to assign particular *cis*-acting regulatory functions to these various regions, transgenes containing different combinations of conserved regions were constructed and analyzed in transient assays.

Mapping of an embryonic activation region

Deletion of the intronic sequences did not affect the early activation of the Ns-E construct in the six β-gal positive embryos obtained (Δ intron; Figure 1B and D). In the two embryos which were recovered at E11, expression was detected in the distal posterior hindlimb. Other late-developing components of the pattern were not influenced and, most importantly, no element of the pattern was lost. We also analyzed the regulatory capacity of the short 0.7 kb *Xho*I−*Apa*I fragment (Xh-A; Figure 1), which contains the transcription initiation site of the mouse gene as well as some upstream sequences. None of the four transgenic embryos that expressed the Xh-A construct showed any of the characteristic patterns associated with the Ns-E construct (Figure 1D). The proportion of β-gal positive embryos among the total transgenic specimens (four out of seven) suggested, however, that a basal promoter activity is located in this fragment. In all four embryos, the β-gal positive structures were located more anteriorly than the s27 limit and were significantly less stained when compared with the Ns-E construct.

The deletion of the 3′-located *Not*I−*Eco*RI fragment (Ns-No, Figure 1B) strongly affected the transgene expression in those structures which were positive in both the Ns-E and the Xh-E constructs (see Figure 1D). This construct was expressed in the trunk of one embryo from pv31 to a more posterior position in the spinal cord and in dorsal root ganglia as well as from s45 in the tail (Figure 4B). In another embryo, the ventral surface of the tailbud was positive from approximately s45. In the four β-gal positive embryos, however, the pv27−30 domain was negative in all tissues (Figure 4B and F and data not shown). In addition to this abnormal posterior expression in the trunk, the three β-gal positive embryos that were 11 days old or older showed strong expression in the anterior two-thirds of the genital tubercle (Figure 4F). The four specimens showed expression in the limb buds, with proximal-posterior restriction, as well as in the umbilical and peri-umbilical abdominal domains, as observed with the Ns-E construct. The absence of proper activation in the early domains suggested that the Ns-No construct lacked regulatory elements required for such a function. Such elements were thus expected to be found in the 3′ part of the locus. We therefore produced further deletions within this region.

The Ns-B construct was not able to activate expression in the trunk (Figures 1B and D, and 4C) and thus was not expressed in any of the domains positive for both the Ns-E and Xh-E constructs. In three of the four positive embryos, expression was found in both the umbilical and abdominal mesoderm. Moreover, in two embryos, the proximal-posterior hindlimb domain was observed. The intensity of staining in the abdominal and hindlimb mesoderm appeared weaker than usual. The late and posterior activation in the spinal cord, dorsal root ganglia and tail mesoderm was absent and no expression was seen in the genital eminence. These unexpected differences between the expression patterns of the Ns-B and Ns-No (shorter) constructs suggested that a negative regulatory element is present between the *Not*I and *Bam*HI sites (see Figure 1B).

We then further deleted the terminal *Xba*I−*Eco*RI fragment to produce the Ns-Xb construct which lacked most of region VIII and region IX (Figures 1B and D, and 5C). None of the four β-gal positive embryos obtained showed either the late or the early patterns as defined by the Ns-E construct (Figure 5C). The absence of the late pattern was even more complete than in the case of the Ns-B construct as a weak expression was found in the forelimb in a single case only. In the other cases the expression patterns were unrelated to any of the Ns-E domains. The region deleted in the Ns-Xb construct contains two clusters of conserved sequences, regions VIII and IX. This result raised the possibility that an important activating function would be achieved by one or both of these regions. An additional deletion was consequently generated which separated the two potential regulatory elements (Ns-H, Figure 1B and D). This construct was equivalent to the Ns-E construct, but lacked region IX, the last conserved DNA domain.

Ten β-gal positive embryos were obtained with the Ns-H construct (Figure 1B and D) and all expressed the transgene in the *Hoxd-11* embryonic domain. Seven specimens showed spatially restricted expression in the trunk. In the six cases where the fetuses were at E10 or older, the anterior limit of expression was found at the level of s25 in axial mesoderm and in the central and peripheral nervous systems (Figure 4H). At E9, an even more anterior limit was found in the trunk mesoderm (Figure 4I), while the limit in the neural tube was still at s25. Even though in this embryo, the s22−23 anterior limit in the trunk was clearly above the hindlimb field, the expression in limb buds of five older fetuses was similar to that seen with the Ns-E transgenic families (Figure 4G and H). In four of these fetuses, the proper pattern was recovered in the genital bud as well. Expression in the fetal limbs and abdominal domains was also detected in the three embryos which were at E10 or older. Consequently, it appeared that the 280 bp *Xba*I−*Hin*dIII fragment was able to restore the expression in the entire embryonic domain, indicating that it contains a major activating function. This same short fragment was also able to rescue expression in the fetal domain. Interestingly, however, the trunk expression domain of the transgene had shifted towards a more anterior position in comparison with the full Ns-E construct, suggesting that part of the regulatory elements that define the precise anterior limit of transgene activation had been lost. As two domains of high interspecies sequence homology were found in the 0.9 kb *Xba*I−*Eco*RI 3′ region (regions VIII and IX, see Figures 1 and 5), one of which was located 3′ of the *Hin*dIII restriction site, we analyzed their potential to act together as a regulatory unit.

Interspecies conservation of the control of Hoxd-11 early activation

Figure 5A shows the alignment of mouse and chick DNA sequences in regions VIII (140 bp) and IX (120 bp). Only very short stretches of homology were found outside these

two domains. The distance between the two domains, however, was strikingly conserved (170 bp for mouse and 169 bp for chick). We used the fact that the *Xba*I restriction site, which is located at the beginning of region VIII, is conserved in both species, to complement the deleted (inactive) mouse Ns-Xb transgene with the 550 bp chicken *Xba*I−*Apa*I genomic fragment. This chimeric mouse−chick transgene gave five β-gal positive fetuses (Mo/Ck, Figure 1B and D). Four out of five showed completely restored patterns in the early and late domains of expression (Figure 5C). Histological examination revealed that distribution of β-gal was exactly the same as that seen for the Ns-E transgene in both the tail and the trunk, including the sharp discrimination between the ureteric bud and metanephric anlage, as well as in the hindgut endoderm and mesoderm (not shown). The chicken 550 bp *Xba*I−*Apa*I *Hoxd-11* fragment could thus substitute for the mouse 3′ flanking fragment to control an early activation function correctly. This demonstrated that regions VIII and IX are indeed solely responsible for this activation, being presumably parts of a composite regulatory unit.

Isolation of a Hoxd-11 embryonic enhancer

To investigate further the regulatory potentials of the 3′ regulatory regions VIII and IX, three additional reporter genes were constructed using a heterologous promoter. The 330 bp *Pst*I−*Hin*dIII fragment, which contains region VIII, was introduced in reverse orientation in front of the *hsp*(3)/*lacZ*(A) enhancer detector gene construct (P-H/hs, Figure 1C and D). Six fetuses showed expression in the posterior trunk region. β-gal positive cells were observed as anteriorly as s16. The intensity of staining increased gradually towards more posterior regions (Figure 4J), reaching a peak level in the regressing primitive streak, the tail bud and the non-segmented mesoderm region. In the three cases where the hindlimb buds were already present, expression was seen in the entire bud and followed the gradual increase in staining, in parallel with that seen in the trunk of these embryos (Figure 4K). In older embryos, low levels of expression was seen in the anterior parts of the *Hoxd-11* early trunk expression domain but nothing in hindlimb buds (not shown), suggesting that the detected β-gal activity in the hindlimbs simply reflected the earlier activation in the trunk rather than the proper mode of *Hoxd-11* Ns-E transgene activation. In the P-P/hs construct (Figure 1C and D), the 400 bp *Hin*dIII−*Pst*I fragment containing region IX was added to the P-H/hs construct. The entire 730 bp *Pst*I fragment which contained both regions VIII and IX was thus present. In this case, the gene was expressed in the tail bud and the trunk, posterior to pv27, but expression was detected neither in more anterior domains nor in the limbs (Figure 4L).

The *Hin*dIII−*Pst*I fragment (P-P/hs) therefore seemed to exert a negative regulatory influence over the activating regulatory region VIII. Moreover, it did not show any embryonic activating capacity when tested alone (H-P/hs). Five transgenic embryos were obtained with the H-P/hs construct (Figure 1C and D), two of which expressed β-gal in the ventral neural tube, in a pattern often observed with the *hsp*(3)/*lacZ*(A) enhancer detector construction alone (J.Zákány, unpublished). One embryo also expressed the transgene in the lung primordia but none of the five expressed the transgene in the domain of the *Hoxd-11* gene.

Discussion

The Hoxd-11 pattern in transgenic mice

In an attempt to characterize the molecular mechanisms that control the spatially and temporally specific activation of *Hox* genes, we have studied the regulation of the *Hoxd-11* gene in transgenic mice. The largest (8 kb Ns-E) fragment used in this study was capable of reproducing some important characteristics of *Hoxd-11* mRNA distribution. The activation of the transgene took place in the posterior end of the 22−23 somite stage embryo, approximately at the level of the future s27. Neuro-ectoderm, pre-somitic and ventral mesoderm, as well as tailgut endoderm, expressed the transgene. This expression pattern was most likely due to an activation of the transgene within the regressing primitive streak. At this stage of delayed gastrulation (in the E9 mouse embryo), the primitive streak is known to contribute to all three germ layers of the posterior body region (Tam and Beddington, 1987). The *Hoxd-11* transgene was thus able to respond to the early activation phase during gastrulation in the primitive streak, a phase during which activation of most *Hox* genes takes place (Gaunt, 1991). The time at which the expression of the transgene was activated corresponds to the occurrence of somitogenesis at the most anterior level at which the expression of the endogenous *Hoxd-11* gene was observed in axial mesoderm (Izpisúa-Belmonte *et al.*, 1991b). In the neural tube, the anterior limit of expression of the transgene was at the same level. This is in contrast to the pattern of *Hoxd-11* specific RNA accumulation whose anterior limit had been previously found at pv25 (Dollé *et al.*, 1991a).

In limb buds, activation of the transgene was seen in the mesoderm. This activation, however, occurred too late. In early hindlimb buds, the β-gal expression was reminiscent of the endogenous *Hoxd-11* specific RNA accumulation. At E11, the pattern started to diverge, with the transgene product being absent from the most distal part, a region that strongly expresses *Hoxd-11* (Dollé *et al.*, 1989). In marked contrast to the normal situation, no expression was seen in the forelimb buds at an early stage. A few hours later, however, a weak activity was reproducibly detected in the posterior-proximal part of the limb. Even though this forelimb pattern was not directly comparable with that of *Hoxd-11*, it was clearly reminiscent of the general posterior specificity of *HoxD* gene expression in the limbs (Izpisúa-Belmonte and Duboule, 1992). Altogether, these results showed that a regulatory region exists around the *Hoxd-11* gene which is capable of activating gene expression in the newly developing posterior body region. The expression of the transgene was observed in all structures which normally express *Hoxd-11*, though with slightly delayed activation when compared with the endogenous gene. However, the pattern of expression in the axial mesoderm was essentially correct.

Genomic fragments that could mimic *Hox* gene activation in the primitive streak at earlier stages were reported from the *Hoxa-7* (Püschel *et al.*, 1991), *Hoxb-1* (Marshall *et al.*, 1992), *Hoxb-4* (Whiting *et al.*, 1991) and *Hoxc-8* (Bieberich *et al.*, 1990) genes previously. As the *Hoxd-11* specific early activation in the primitive streak and hindlimb buds was rather faithfully reproduced with the 5 kb Xh-E construct, we set out to localize the responsible *cis*-acting element. We used interspecific DNA sequence comparison to identify

conserved genomic regions (Renucci *et al.*, 1992) that could be functionally assayed by additional deletions, choosing two distantly related species.

Definition of an early activating region

Seven domains of high interspecies sequence homology were found in non-protein-coding regions of the *Hoxd-11* locus. Five of these overlapped with, or were included in, the mRNA, whereas two were found in 3′ flanking sequences. The analysis of various constructs which contained different combinations of these domains allowed us to evaluate their regulatory potentials. The presence or absence of intronic sequences did not appreciably modify the early expression pattern, suggesting that region IV is not required for the early activation. Similarly, the presence of the conserved regions I and II in the short Xh-A construct, which contained the *Hoxd-11* start site and upstream region, was not sufficient to recapitulate any specific character of the endogenous expression pattern. The variable expression patterns obtained with this short fragment were most likely due to regulatory influences conferred by the different integration sites.

In addition to regions I and II, regions III−V and III−VII were present in the Ns-B and Ns-Xb constructs, respectively. None of the above combinations, however, was able to mimic the early activation pattern. It thus appears that the early activation of the transgene is dependent upon the presence of the 3′ terminal 0.9 kb fragment. This fragment contains regions VIII and IX and was referred to as the early activation region. Some transgenes could nevertheless exhibit significant activities in the *Hoxd-11* late expression domains without this early activating region. This suggests the presence of additional positively as well as negatively acting regulatory elements.

Late-acting positive and negative regulatory regions

The 3′ deletions have established the existence of a second, late-acting, positive regulatory region. The Ns-No construct contained the *Hoxd-11* transcription initiation site (regions I and II and part of region III). This construct was activated in the late pattern, in contrast to the Xh-E construct. This suggests that the activation in the late domains relies on the presence of the 5′-located *Nsi*I−*Xho*I fragment. The *Hoxd-11* transgene expression could thus be initially controlled by the 3′ regulatory region, in the primitive streak, pre-somitic mesoderm and somites even in the absence of the 5′ regulatory region. Later, the presence of the 5′ and 3′ regulatory sequences may be required simultaneously to guarantee a correct expression in the structures which derive from these domains, e.g. in the pre-vertebrae. The activity of the *Hoxd-11* promoter, at later stages, may therefore result from a cooperative interaction between two distantly located activation regions.

The expression patterns produced by the Ns-No, Ns-B and Ns-Xb constructs revealed the existence of several negatively acting regions. In the genital eminence, for example, expression was observed with the Ns-No construct whereas the longer Ns-B test gene was negative. Furthermore, most of the other Ns-No-associated expression domains were significantly reduced, either in the Ns-B expressing embryos or, even more, with the Ns-Xb construct. This suggested that a rather large negatively acting regulatory region is present in the *Not*I−*Xba*I fragment, i.e. within the transcriptional unit, around regions IV−VII. The importance of negative regulatory regions in specifying the anterior limit of transgene expression in late fetal tissues has been previously reported for the murine *Hoxa-7* (Püschel *et al.*, 1990) and *Hoxc-8* (Bieberich *et al.*, 1990) genes.

The 3′ activating region is a bipartite element

The expression of the inactive Ns-Xb construct could be rescued by the addition of the adjacent 270 bp *Xba*I−*Hin*dIII fragment (Ns-H), which thus identified region VIII as the early activating element. While the expression was restored in the entire early domain as well as in much of the late domain, the spatial restriction of transgene expression during the early phase of activation was shifted anteriorly relative to the full Ns-E construct (Figure 6). This suggested that the correct spatial activation involves another regulatory function, located further 3′, that would cooperate with region VIII to define the proper anterior limit during the early activation. This region was further studied by the production of a chimeric transgene which replaced the 3′-located murine *Xba*I−*Eco*RI fragment with a corresponding fragment of the chicken *Hoxd-11* gene. The nucleotide sequence of this chicken fragment shows high levels of homology with its mouse counterpart only in regions VIII (140 bp, 87% identity) and IX (120 bp, 89% identity). Interestingly, the distance between these two regions is strikingly conserved (186 bp in mouse and 185 bp in chick). The chimeric construct reproduced an original activation pattern similar

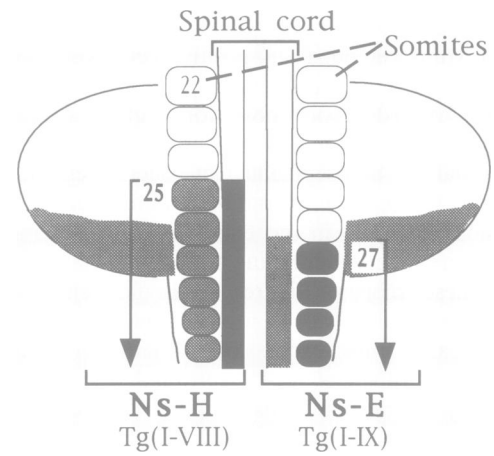


Fig. 6. Cooperations between the enhancer (region VIII) and the negative element (region IX) in the establishment of the *Hoxd-11* expression specificity. A schematic representation of the posterior body region of an E10 embryo is shown, with the somites and the corresponding levels of the spinal cord. Shading of these structures indicates β-gal activity observed with either the Ns-H (from s25, on the left-hand side) or the Ns-E (from s27, on the right-hand side) constructs. The expression domains extend posteriorly (arrows). The Ns-H construct contains the enhancer (Tg I−VIII) while the Ns-E construct contains, in addition, the negative element (Tg I−IX). The spinal cord on the left and the somites on the right are shown in black to indicate that these domains of expression correspond to those of the endogenous *Hoxd-11* gene. The somites on the left and the spinal cord on the right are stippled to show that these expression domains are abnormal in their extent. These deviations from the activity of the endogenous *Hoxd-11* gene may reflect the capacity of the negative element to prevent enhancer-mediated expression in the somites anterior to s27 as well as a requirement for the functional isolation of these two cooperating elements to reach the correct level in the spinal cord. This functional isolation cannot be reproduced in transgenic mice suggesting that it depends on a *cis*-acting control from a *HoxD* region that is not present in the *Nsi*I−*Eco*RI fragment.

to that observed with the mouse Ns-E construct. This functional equivalence, in the transgenic mouse assay, of the mouse and chicken fragments, added to the modifications of regulatory specificity following the removal of region IX (the Ns-H construct), strongly suggests that regions VIII and IX represent major *Hoxd-11* regulatory elements. It also indicates that the two elements cooperate in setting up the proper early activation of the *Hoxd-11* expression.

When assayed upstream of the *hsp68* promoter, region VIII showed some features of transcriptional enhancers. Interestingly, it caused selective activation in the tail bud but the expression domain extended abnormally far towards the anterior, even more so than with the Ns-H construct. In contrast, the *hsp68* promoter construct that contained the original combination (regions VIII and IX) was not expressed anterior to the level of pv27 in the trunk. These observations confirm the presence of a negative regulatory function mediated by region IX. This negative control region was capable of preventing the activation of the transgene in ectopic anterior expression domains, probably by negatively controlling a premature activity of the enhancer element VIII. Heat-shock promoter experiments also showed that, although the enhancer was sufficient for activation in the trunk, it could not direct specific activation in the hindlimbs. This is in contrast with the observed ability of the same element to rescue limb bud expression in its original context.

Regions VIII and IX contain eight conserved sequence motifs somehow related to the consensus binding site for nuclear hormone receptors. Some of these motifs are found in arrangements that have previously been shown to bind thyroid hormone receptor and RAR (Umesono *et al.*, 1988), or RAR-RXR heterodimers (Yu *et al.*, 1991; Durand *et al.*, 1992), and regressing primitive streak expresses high levels of some members of the nuclear receptor family (Ruberte *et al.*, 1991).

Two modes of activation

The cooperation of regions VIII and IX may play a crucial role in restricting *Hoxd-11* promoter activity along the verebral column. When region IX is removed from the Ns-E construct (giving Ns-H), the anterior limit of expression in the column is shifted from the sacral region (s27 in Ns-E) towards the lower lumbar region (s25). In absence of the negative element IX, the enhancer, together with other negatively acting element(s) located upstream, may thus activate with a proper anterior limit (s25) in the primitive streak and the neural tube. However, expression in the presomitic mesoderm was too anterior, suggesting that region IX is required to properly set the posterior expression level in somites.

These results suggest that two distinct mechanisms may operate in the control of *Hoxd-11* activation by regions VIII and IX. In the mesoderm, proper regulation may rely on a close functional association of the two halves of this bipartite regulatory region (inactive in s25−27). In contrast, proper expression in both the primitive streak and neural tube apparently requires the functional isolation of region IX from region VIII (active in s25−27). Simultaneous requirements for functional isolation and association of regulatory sequences may provide an explanation to the almost complete conservation of the distance between these two regulatory elements in both the avian and mammalian *HoxD* complexes.

Materials and methods

Recombinant DNA techniques

For the pGemE/E subclone, the 8 kb *Eco*RI fragment containing the *Hoxd-11* gene was isolated from cosB (Izpisúa-Belmonte *et al.*, 1991b) and cloned into the *Eco*RI site of pGem-7Zf (Promega). For the Ns-E construct, the *lacZ* gene (*Pst*I fragment from PMC1871, Pharmacia) was inserted into pPolyIII-D (Lathe *et al.*, 1987), excised with *Not*I and introduced into the *Not*I site of the *Hoxd-11* first exon in the pGemE/E plasmid. A short DNA fragment made of two partially complementary oligonucleotides (5′-CGTTCCGATACCGA-3′ and 5′-GTATCGGAACGTCG-3′) was then inserted into the *Sfi*I site (from the pPolyIII-D linker) to adjust the reading frames of the *lacZ* and *Hoxd-11* coding sequences. For injection the 5.5/*lacZ* fragment was excised with *Nsi*I. For the Ns-E construct containing *lacZ* and SV40 poly(A) sequence, the *lacZ* gene (*Bam*HI fragment from pMC1871) was inserted into the *Bgl*II site of pSG5 (Green *et al.*, 1988) and excised from this plasmid as a *Bam*HI−*Sal*I fragment containing the 140 bp SV40 poly(A) signal. This fragment was ligated with linkers 1 and 2 and inserted in-frame into the *Not*I site of the *Hoxd-11* first exon in pGemE/E. Linker 1 was made of two partially complementary oligonucleotides, 5′-GGCCAGTCGACGCATGCTT-3′ and 5′-GATCAAGCATGCGTCGACT-3′, and linker 2 of 3′-TCGAACTAGTACTGC-5′ and 3′-GGCCGCAGTACTAGT-5′. This gave the pGemE/ElacZpA construct. For injection, this Ns-E construct was excised with *Nsi*I. The Xh-E construct was excised with *Xho*I and *Nsi*I, the Ns-No construct with *Nsi*I and *Not*I, the Ns-B construct with *Nsi*I and *Bam*HI, and the Ns-Xb construct with *Nsi*I and *Xba*I. For the intron deletion construct, the genomic *Not*I−*Acc*I 1130 bp fragment from pGemE/ElacZpA was exchanged with the *Not*I−*Acc*I *Hoxd-11* cDNA (Izpisúa-Belmonte *et al.*, 1991b) fragment (390 bp). This generated a transgene with its intronic sequences deleted. pGemE/ElacZpA was digested with *Xba*I and *Not*I, and the fragments *Xba*I−*Not*I (7.5 kb), *Not*I−*Xba*I (2.5 kb) and *Xba*I (3.9 kb, containing pGem7) were purified. The *Acc*I−*Xba*I 1.5 kb fragment from *Not*I−*Xba*I (2.5 kb) was ligated with the 290 bp *Not*I−*Acc*I *Hoxd-11* cDNA fragment. The resulting *Not*I−*Xba*I (1.8 kb) fragment was subcloned for amplification and further ligated with the *Xba*I−*Not*I 7.5 kb fragment. The resulting 9.3 kb *Xba*I fragment was cloned into the *Xba*I 3.9 kb fragment. Clones were selected for proper orientation. For injection, the fragment was excised with *Nsi*I. For the XhA construct, the 1.1 kb *Xho*I−*Apa*LI fragment from *Hoxd-11* 5′ region was half-filled with Klenow at the *Apa*I site and fused with *lacZ* into the phsp(3)lacZpA plasmid (kindly provided by J.Rossant), linearized with *Sal*I and *Nco*I (half-filled with Klenow at the *Nco*I site). The Xh-A fragment was excised for injection with *Hin*dIII and *Xmn*I. For the Ns-H construct, the 280 bp *Xba*I−*Hin*dIII fragment containing region VIII was cloned in pGem7 (pGem280*Xba*I/*Hin*dIII). The 10 kb *Xba*I−*Xba*I fragment from pGemE/ElacZpA was cloned into the *Xba*I site of pGem280*Xba*I/*Hin*dIII, and clones were selected for proper orientation. The Ns-H construct was excised for injection with *Nsi*I. For the chimeric Mo/Ck construct, the chicken *Hoxd-11* gene was isolated from CosII (Izpisúa-Belmonte *et al.* 1991a) as a 9.5 kb *Bam*HI fragment, which was cloned in pBluescriptII KS (giving pKS9.5). *Hoxd-11* regions VIII and IX are located in the *Not*I−*Bam*HI 1.25 kb fragment of pKS9.5, which was cloned into the *Not*I/*Bam*HI sites of pBluescriptIIKS (pKS1.25). The chicken *Xba*I−*Apa*I 550 bp genomic fragment was excised from pKS1.25 and cloned into the *Xba*I/*Apa*I sites of pBluescriptIIKS, then excised as a *Xba*I−*Kpn*I fragment and cloned into the *Xba*I/*Kpn*I sites of pGem7, resulting in the pGem550 *Xba*I/*Apa*I construct. The 10 kb *Xba*I−*Xba*I fragment from pGemE/ElacZpA was cloned into the *Xba*I site of the pGem550 *Xba*I/*Apa*I construct. Clones were selected for proper orientation and the Mo/Ck fragment was excised with *Nsi*I. For the P-H/hs, H-P/hs and P-P/hs constructs, the 730 bp *Pst*I−*Pst*I fragment containing regions VIII and IX was cloned into the *Pst*I site of pSP72 (Promega). Both orientations were selected and fragments (orientations 1 and 2) were excised with *Sal*I and *Sph*I and further cloned into the *Sph*I and *Sal*I sites of phsp(3)lacZpA. P-H/hs and P-P/hs fragments were excised with *Hin*dIII and *Xmn*I or *Sph*I and *Xmn*I respectively, while the H-P/hs fragment was excised with *Hin*dIII and *Xmn*I. All fragments were purified from agarose gels by phenol/chloroform extraction or using Geneclean.

Production of transgenic mice, X-gal staining and histology

Transgenic mice were produced according to standard procedures (Hogan *et al.*, 1986) using B6D2F2 eggs. Transgenic families carrying the 5.5/*lacZ* Ns-E transgenes were established by crossing the founder animals to B6D2F1 mates and inter-crossing transgenic (F1) progeny. Families #370, 387, 389, 391 and 397 were kept as homozygous lines. The other constructs were tested in transient assays. Whole mount enzyme detection and histological and transgene detection procedures were carried out as previously described

(Zákány *et al.*, 1988). Histological sections in Figure 3C, D, F, G, H, K and L and Figure 4E and F were counterstained with eosin.

In the case of the Ns-E 8 kb construct, the cumulative results in Figure 1D are pooled from seven independent insertion events with the 5.5/*lacZ* construct, (transgenic families 397, 389 and 387, plus one transgenic mouse embryo produced in transient assay, all exhibiting the characteristic elements of the pattern; two families, and one transient assay derived embryo, that do not express the transgene in the posterior regions) and eight independent insertion events with the Ns-E construct, isolated from pGemE/ElacZpA. In these last cases, four of them expressed in at least one of the characteristic expression domains. All were observed in transient assays. The cumulative results with the Xh-E transgene are pooled from six independent integration events produced with the 5.5/*lacZ* Xh-E construct, three of which expressed the transgene in the characteristic expression domains, and five independent integration events obtained with the Xh-E construct isolated from the pGemE/ElacZpA, four of which expressed the transgene in the characteristic domain.

The time course was established in detail with family 397 and verified at some stages with families 389 and 387. Besides the recurrent elements of the pattern, a number of expression sites were observed that were specific for individual transgenic families or particular embryos in transient assays. Expression in the head mesoderm of late E10 (Theiler stage 17) or in older embryos of the Ns-E family #397 is one example (see Figure 3E). Such patterns, however, do not show comprehensible linkage to the *Hoxd-11* transgene fragment and some appeared as superimpositions over the otherwise correct *Hoxd-11* specific pattern. Four such additional expression domains out of the 12 cases where proper expression was obtained in the posterior body regions, suggests that the *Nsi*I–*Eco*RI fragment does not contain sufficient regulatory information to prevent anterior expression completely when present at certain integration sites.

Acknowledgements

We would like to thank Hildegard Falkenstein and Dr Juan-Carlos Izpisúa-Belmonte for their help in the initial phase of this work as well as Drs Janet Rossant and Achim Gossler for the phsp(3)lacZpA plasmid, and Dave Stoddard for animal husbandry and expert technical assistance. We also thank the EMBL photolab as well as Gábor Lamm, Armand Renucci and Gilbert Urier for discussions and suggestions on the manuscript. M.G. was on leave from the LGME du CNRS, Unité INSERM 184, Faculté de Médecine, 11 rue Humann, 67085 Strasbourg Cedex, France. J.Z. was on leave from the Institutes of Biphysics and Plant Physiology of the Biological Research Centre of the Hungarian Academy of Sciences, Szeged, Hungary. This work was supported by funds from the EMBL and HFSPO.

References

Bieberich,C.J., Utset,M.F., Awgulewitsch,A. and Ruddle,F.H. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 8462–8466.
Chisaka,O. and Capecchi,M.R. (1991) *Nature*, **350**, 473–479.
Chisaka,O., Musci,T. and Capecci,M. (1992) *Nature*, **355**, 516–520.
Dollé,P., Izpisúa-Belmonte,J.-C., Falkenstein,H., Renucci,A. and Duboule,D. (1989) *Nature*, **342**, 767–772.
Dollé,P., Izpisúa-Belmonte,J.-C., Boncinelli,E. and Duboule,D. (1991a) *Mech. Dev.*, **36**, 3–14.
Dollé,P., Izpisúa-Belmonte,J.-C., Brown,J.M., Tickle,C. and Duboule,D. (1991b) *Genes Dev.*, **5**, 1767–1776.
Duboule,D. (1992) *BioEssays*, **14**, 375–384.
Duboule,D. and Dollé,P. (1989) *EMBO J.*, **8**, 1497–1505.
Durand,B., Saunders,M., Leroy,P., Leid,M. and Chambon,P. (1992) *Cell*, **71**, 73–85.
Gaunt,S.J. (1991) *BioEssays*, **13**, 234–242.
Gaunt,S.J., Sharpe,P.T. and Duboule,D. (1988) *Development* (suppl.), **104**, 169–179.
Graham,A., Papalopulu,N. and Krumlauf,R. (1989) *Cell*, **57**, 367–378.
Green,S., Issemann,I. and Sheer,E. (1988) *Nucleic Acids Res.*, **16**, 369.
Hogan,B., Costantini,F. and Lacy,E. (1986) *Manipulating the Mouse Embryo*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Izpisúa-Belmonte,J.-C. and Duboule,D. (1992) *Dev. Biol.*, **152**, 26–36.
Izpisúa-Belmonte,J.-C., Dollé,P., Renucci,A., Zappavigna,V., Falkenstein,H. and Duboule,D. (1990) *Development*, **110**, 733–746.
Izpisúa-Belmonte,J.-C., Tickle,C., Dollé,P., Wolpert,L. and Duboule,D. (1991a) *Nature*, **350**, 585–589.
Izpisúa-Belmonte,J.-C., Falkenstein,H., Dollé,P., Renucci,A. and Duboule,D. (1991b) *EMBO J.*, **10**, 2279–2289.
Jegalian,B.G. and De Robertis,E.M. (1992) *Cell*, **71**, 901–910.
Kessel,M., Balling,R. and Gruss,P. (1990) *Cell*, **61**, 301–308.
Lathe,R., Violette,J.L. and Clark,A.J. (1987) *Gene*, **57**, 193–201.
LeMouellic,H., Lallemand,Y. and Brûlet,P. (1992) *Cell*, **69**, 251–264.
Lufkin,T., Dierich,A., LeMeur,M., Mark,M. and Chambon,P. (1991) *Cell*, **66**, 1105–1119.
Lufkin,T., Mark,M., Hart,C.P., Dollé,P., LeMeur,M. and Chambon,P. (1992) *Nature*, **359**, 835–841.
McGinnis,W. and Krumlauf,R. (1992) *Cell*, **68**, 283–302.
Marshall,H., Nonchev,S., Sham,M.H., Muchamore,I., Lumsden,A. and Krumlauf,R. (1992) *Nature*, **360**, 737–741.
Morgan,B.A., Izpisúa-Belmonte,J.-C., Duboule,D. and Tabin,C.J. (1992) *Nature*, **358**, 236–239.
Pollock,R.A., Jay,G. and Bieberich,C.J. (1992) *Cell*, **71**, 911–923.
Püschel,A.W., Balling,R. and Gruss,P. (1991) *Development*, **112**, 279–284.
Renucci,A., Zappavigna,V., Zákány,J., Izpisúa-Belmonte,J.-C., Bürki,K. and Duboule,D. (1992) *EMBO J.*, **11**, 1459–1468.
Rogina,B., Coelho,C.N.D., Kosher,R.A. and Upholt,W.B. (1992) *Dev. Dynamics*, **193**, 92–101.
Ruberte,E., Dollé,P., Chambon,P. and Morris-Kay,G. (1991) *Development*, **111**, 45–60
Sham,M.H., Hunt,P., Nonchev,S., Papalopulu,N., Graham,A., Boncinelli,E. and Krumlauf,R. (1992) *EMBO J.*, **11**, 1825–1836.
Scott,M. (1992) *Cell*, **71**, 551–553.
Tam,P.P.L. and Beddington,R.S.P. (1987) *Development*, **99**, 109–126.
Umesono,K., Giguere,V., Glass,C.K., Rosenfeld,M.G. and Evans,R.M. (1988) *Nature*, **336**, 262–265..
Whiting,J., Marshall,H., Cook,M., Krumlauf,R., Rigby,P.W.J., Scott,D. and Allemann,R.K. (1991) *Genes Dev.*, **5**, 2048–2059.
Yokouchi,Y., Sasaki,H. and Kuroiwa,A. (1991) *Nature*, **353**, 443–445.
Yu,V.C., Delsect,C., Anderson,B., Holloway,J.M., Devary,O., Naar,A.M., Kim,S.Y., Bontin,J.-M., Glass,C.K. and Rosenfeld,M.D. (1991) *Cell*, **67**, 1251–1266.
Zákány,J., Tuggle,C.K., Patel,M. and Nguyen-Huu,M.C. (1988) *Neuron*, **1**, 679–691.

Received on April 29, 1993; revised on May 26, 1993