Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 9339–9345, September 1996
Colloquium Paper

This paper was presented at a colloquium entitled "Biology of Developmental Transcription Control," organized by Eric H. Davidson, Roy J. Britten, and Gary Felsenfeld, held October 26–28, 1995, at the National Academy of Sciences in Irvine, CA.

The conserved role of *Krox-20* in directing *Hox* gene expression during vertebrate hindbrain segmentation

STEFAN NONCHEV*, MARK MACONOCHIE*, CHRISTINE VESQUE†, SAM APARICIO‡, LINDA ARIZA-MCNAUGHTON*, MIGUEL MANZANARES*, KAMALA MARUTHAINAR*, ATSUSHI KUROIWA§, SYDNEY BRENNER‡, PATRICK CHARNAY†, AND ROBB KRUMLAUF*¶

*Laboratory of Developmental Neurobiology, Medical Research Council, National Institute for Medical Research, The Ridgeway, Mill Hill, London, NW7 1AA, United Kingdom; †Unite 368 de L'Institut National de la Santé et de la Recherche Médicale, Ecole Normale Superieure, 46 rue d'Ulm, 75230 Paris, Cedex 05, France; ‡Molecular Genetics Unit, Department of Medicine, Addenbrookes Hospital, Cambridge, CB2 2QQ, United Kingdom; and §Department of Molecular Biology, School of Science, Nagoya University, Furo-cho, Chilusa-ku, Nagoya 464–01, Japan

ABSTRACT **Transient segmentation in the hindbrain is a fundamental morphogenetic phenomenon in the vertebrate embryo, and the restricted expression of subsets of *Hox* genes in the developing rhombomeric units and their derivatives is linked with regional specification. Here we show that patterning of the vertebrate hindbrain involves the direct upregulation of the chicken and pufferfish group 2 paralogous genes, *Hoxb-2* and *Hoxa-2*, in rhombomeres 3 and 5 (r3 and r5) by the zinc finger gene *Krox-20*. We identified evolutionarily conserved r3/r5 enhancers that contain high affinity Krox-20 binding sites capable of mediating transactivation by Krox-20. In addition to conservation of binding sites critical for Krox-20 activity in the chicken *Hoxa-2* and pufferfish *Hoxb-2* genes, the r3/r5 enhancers are also characterized by the presence of a number of identical motifs likely to be involved in cooperative interactions with Krox-20 during the process of hindbrain patterning in vertebrates.**

The developing central nervous system in vertebrates plays an important role in head morphogenesis because it generates cranial neural crest, which gives rise to the bone and connective tissue of the face (1). In the hindbrain, axial patterning is achieved through a process of segmentation, in which the developing metameric units (termed rhombomeres) form lineage-restricted cellular compartments (2–4). Grafting experiments in avian embryos have demonstrated that the rhombomeres and their cranial neural crest derivatives have a prepattern and will generate structures typical of their anterioposterior origin in a cell-autonomous manner when placed in ectopic locations (5, 6). On the basis of expression patterns and phenotypes in gene disruption experiments in the mouse, the family of *Hox* homeobox genes are widely believed to play a fundamental role in regulating the anterioposterior identity of the rhombomeres and neural crest (reviewed in refs. 7–11). While *Hox* genes are an integral part of the process of specifying regional variation in the developing hindbrain, little is known about the cascade of events that regulate their expression or segmentation itself. Therefore, considerable effort has been placed upon identifying the upstream regulatory factors that impose segment-restricted expression of the *Hox* genes as a means of beginning to unravel the molecular mechanisms underlying the generation of segments and their axial specification.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

With respect to this question, transgenic and mutational studies have recently shown that the zinc finger gene *Krox-20* is an essential component in controlling hindbrain segmentation. *Krox-20* is normally expressed in rhombomeres 3 and 5 (r3 and r5), and in embryos with homozygous null alleles of the gene, these segments partially form but are not maintained in later stages of development (12, 13). Furthermore, *Krox-20* has been shown to directly regulate the *Hoxb-2* and *Hoxa-2* genes in r3 and r5, through cis-acting sequences in the 5′-flanking regions of the genes (14, 15). While the Krox-20 binding sites in the enhancers of these two *Hox* genes are necessary for regulatory activity, they are not sufficient to direct expression of a *lacZ* reporter in transgenic mice, suggesting that additional factors are required.

The process of hindbrain segmentation appears to be evolutionarily conserved in vertebrates because the number of rhombomeres, the neuroanatomical organization and the patterns of expression of many genes, including transcription factors, tyrosine kinase receptors, and signaling molecules, also display a high degree of conservation (reviewed in ref. 2). Use of transgenic mice to functionally test regulatory regions of *Hox* genes, combined with sequence comparisons, has proved to be very useful in delimiting critical conserved motifs implicated in regulation of segmental expression (16–20). Therefore, we wanted to use a similar strategy of evolutionary comparison with other vertebrates to determine if *Krox-20* had a common role in regulating *Hox* genes in r3 and r5. Furthermore, we hoped that conservation of motifs in addition to Krox-20 binding sites themselves, might be a first step in helping to identify the cofactors required by *Krox-20* for regulatory activity. In this study, we have functionally tested 5′-flanking regions of the chicken *Hoxa-2* and the *Fugu* (pufferfish) *Hoxb-2* genes and performed sequence comparisons with the mouse genes. Our findings show that the general degree of sequence identity in 5′-flanking regions between group 2 paralogs is low, but we detected conserved Krox-20 binding sites and a new associated motif potentially involved in cooperating in segmental regulation. Therefore, we conclude that roles for *Krox-20* in r3 and r5 are conserved in other vertebrates.

MATERIALS AND METHODS

Whole-Mount *in Situ* Hybridization, DNA Constructs, and Transgenic Analysis. Whole-mount *in situ* hybridization of

Abbreviations: r3 and r5, rhombomeres 3 and 5; dpc, days post coitum.
¶To whom reprint requests should be addressed.

mouse embryos was performed as described (21) with use of specific fragments from the *Hoxa-2* and *Hoxb-2* genes to generate a T7 transcribed riboprobe labeled with digoxigenin. A set of fragments 5′ to the chicken *Hoxa-2* ATG were inserted in the vector 1084 (construct 8 of ref. 22) and used as reporter constructs in the transgenic experiments (see Fig. 2). The embryos shown on Fig. 1 *E* and *F* were generated with a 1.2-kb *Eco*RI fragment (Fig. 2, construct 2) and a 0.5-kb *Eco*RI/*Dra*I fragment (Fig. 2, construct 4), respectively. A pufferfish 2.8-kb region encompassing the first exon, the intron, and 1.8 kb 5′ to the ATG (Fig. 2, construct 6) was used to obtain the embryo shown in Fig. 1*G*. Transgenic embryos were generated by microinjection of fertilized eggs from crosses between F_1 hybrids (CBA × C57). PCR analysis and β-galactosidase staining of embryos were performed as described (14, 17, 18, 22).

Sequencing of the Chicken *Hoxa-2* and *Fugu Hoxb-2* r3/r5 Enhancers. Both strands of all regulatory regions were sequenced by the dideoxynucleotide method on an Applied Biosystems ABI373A DNA sequencer according to the manufacturer's instructions. Two phages overlapping the pufferfish *Hoxb-2* 5′ region were obtained by walking upstream from the *Hoxb-1* locus. The *Hin*dIII/*Not*I fragment tested represents an end clone from the phage N26-8, which contains 1.8 kb 5′ of the pufferfish *Hoxb-2* ATG.

Comparative Analysis of Regulatory Sequences. The sequence homology search was performed first by pairwise comparisons using the GCG programs BESTFIT and BETTERFIT on both strands. ICATOOLS and PRINTREPEATS programs (23) were then used to search larger tracts of sequence for short similarity blocks, to analyze in more detail the conserved regions, and to generate the multiple nucleotide sequence comparison.

RESULTS

Identification and Mapping of a Chicken *Hoxa-2* r3/r5 Enhancer. In the mouse, the anterior expression boundaries of the group 2 paralogous genes differ. Expression of *Hoxa-2* maps to the r1/r2 junction (8) and *Hoxb-2* to the r2/3 boundary (24); however, within their domains of expression, some specific segments have elevated levels (reviewed in refs. 9 and 11). For both of these genes, there is an up-regulation of expression in r3 and r5, and only in the case of *Hoxb-2* is there additionally strong staining in r4 (Fig. 1 *a* and *b*). The r3 and r5 domains are those we have previously shown to be directly regulated by *Krox-20* (14, 15). This is illustrated by the observation that enhancers from the *Hoxb-2* gene (Fig. 1*c*) and the *Hoxa-2* gene (Fig. 1*d*) will impose r3/r5-restricted expression on a *lacZ* reporter gene when tested in transgenic mice.

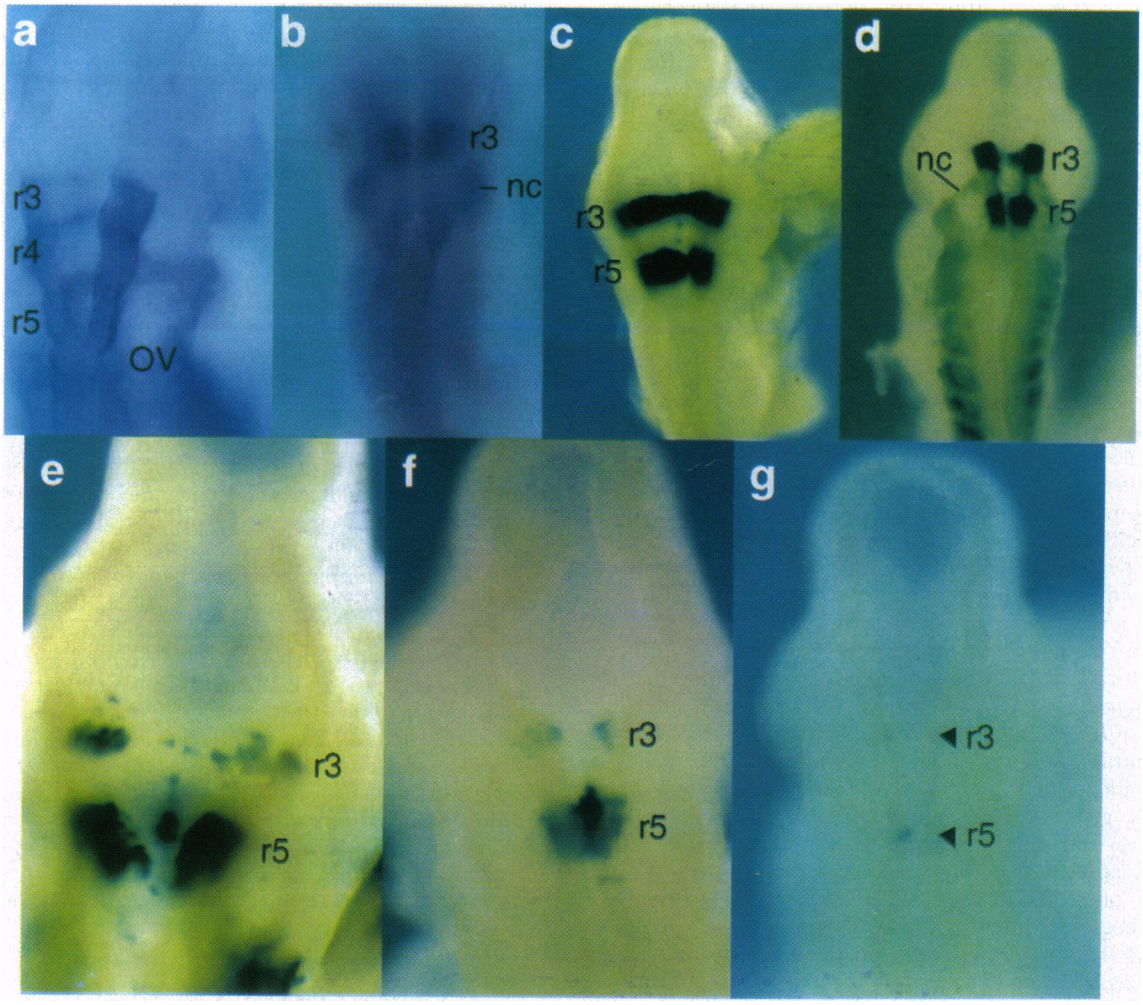

FIG. 1. Analysis of the r3/r5 up-regulation of group 2 *Hox* gene paralogs and homologs in transgenic embryos. (*a* and *b*) Whole-mount *in situ* hybridization of 9.5-days post coitum (dpc) (*a*) and 9.0-dpc (*b*) mouse embryos with *Hoxb-2* and *Hoxa-2* probes, respectively. Note the high levels in r3 and r5. (*c* and *d*) Staining in 9.5-dpc transgenic mouse embryos driven by the *Hoxb-2* and *Hoxa-2* r3/r5 enhancers, respectively. (*e* and *f*) Transgenic mouse embryos (9.5 dpc) produced with a 1.2-kb *Eco*RI (Fig. 2, construct 2) and a 0.5-kb *Eco*RI/*Dra*I (Fig. 2, construct 4) chicken *Hoxa-2* regulatory fragment, respectively. (*g*) Expression in a 9.5-dpc transgenic embryo containing a pufferfish 2.8-kb genomic fragment (Fig. 2, construct 6) encompassing the first exon, the intron, and a 1.8-kb of sequence upstream of the ATG. OV, otic vesicle; nc, neural crest.

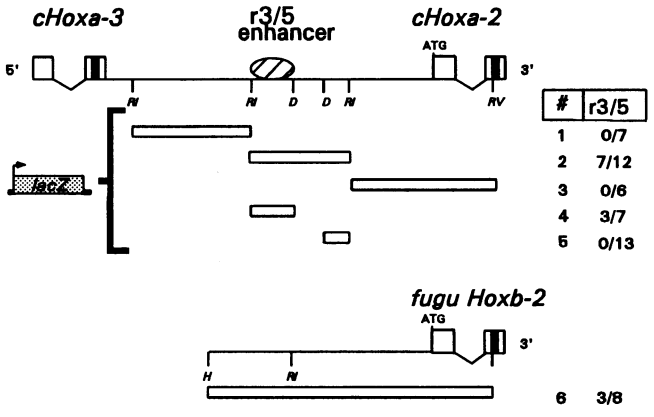

FIG. 2. Mapping of the chicken *Hoxa-2* and pufferfish *Hoxb-2* r3/r5 enhancer elements in transgenic mice. The relative positions of the chicken *Hoxa-2* and *Hoxa-3* and the pufferfish *Hoxb-2* genomic loci are depicted on the left with the restriction fragments used in transgenic analysis. At the right, the construct number (#) and the fraction of embryos expressing the reporter lacZ gene in r3 and r5 are indicated. RI, *Eco*RI; D, *Dra*I; RV, *Eco*RV; H, *Hin*dIII.

The chicken *Hoxa-2* gene has a pattern of expression similar to that of the mouse with an anterior boundary mapping to r1/r2; however, the levels of expression in r3 and r5 were not obviously higher (25). This raised the question as to whether the chicken *Hoxa-2* gene was controlled by *Krox-20*. Therefore, we subcloned the chicken *Hoxa-2* gene and its 5′-flanking region extending to the next upstream gene (*Hoxa-3*) and mapped the relative positions of their initiation ATGs (Fig. 2). Because the r3/r5 enhancer of the mouse gene is located upstream, we first looked for potential *Krox-20* binding activity in this region. Bacterially expressed Krox-20 protein was used in *in vitro* analysis to immunoprecipitate fragments in the *Hoxa-2/Hoxa-3* intergenic region. We found that a 1.2-kb *Eco*RI fragment (Fig. 2) was preferentially bound by the protein (data not shown).

We also generated a series of constructs linking the 5′-flanking regions of the chicken gene to a *lacZ* reporter gene under the control of a basal heterologous promoter (Fig. 2). The functional activity of these constructs (1–5) was tested in transgenic analysis. Constructs 1 (0/7) and 3 (0/6) had no activity in transgenic embryos, whereas construct 2, carrying a 1.2-kb *Eco*RI fragment, generated specific transgene staining in r3 and r5 (Fig. 1*e*). In the majority of the embryos generated with construct 2, the expression in r5 was at high levels and uniformly distributed throughout the body of the rhombomere. In contrast, r3 displays patchy groups of positive cells either dispersed in multiple foci in the rhombomere or specifically localized and confined to its boundaries (Fig. 1*e*). The difference in r3 staining is not due to a temporal alteration, as embryos harvested at earlier and later times display a similar pattern in r3. Therefore, the overall staining pattern is very similar to that obtained with the mouse enhancer (Fig. 1*d*) with the exception of the lower level in r3, which appears to be a genuine difference in the ability of the chicken elements to function in mice.

The immunoprecipitation and transgenic results with the 1.2-kb *Eco*RI fragment suggest that *Krox-20* is involved in regulating the chicken *Hoxa-2* gene in r3 and r5. Further deletion constructs were tested to narrow down the minimal enhancer elements with r3 and r5 activity. Construct 4 (3/7) containing an 0.5-kb *Eco*RI–*Dra*I subfragment was able to mediate r3 and r5 expression (Fig. 1*f*). The pattern of expression stimulated by this fragment was similar to that with the entire 1.2-kb enhancer. However, the number of positive embryos expressing this pattern was generally lower compared with the 1.2-kb enhancer, as were the relative levels of staining in rhombomeres of positive embryos. This suggests that regions modulating the relative efficiency of the r3/r5 enhancer might have been deleted in the minimal *Eco*RI–*Dra*I construct.

An r3/r5 Enhancer in the *Fugu Hoxb-2* 5′-Flanking Region. Based on the success in identifying a conserved enhancer in the *Hoxa-2* gene from chicken, we wondered if similar regulatory regions are conserved in the *Hoxb-2* gene and if this conservation could extend over a greater evolutionary distance. From initial *in situ* hybridization analysis of the *Hoxb-2* gene in zebrafish embryos, the boundary of expression appears to map to the same r2/r3 boundary as in the mouse (data not shown). Again, there is no apparent up-regulation of *Hoxb-2* expression in r3 and r5 that could obviously be used to implicate *Krox-20* in its control, despite the fact that *Krox-20* is indeed expressed in r3 and r5 in zebrafish hindbrain (26). Therefore, we used a similar approach to that described above for the chicken *Hoxa-2* gene and tested a 1.8-kb 5′-flanking region of the *Fugu* (pufferfish) *Hoxb-2* gene for regulatory activity in transgenic mice. We used the pufferfish *Hoxb-2* gene instead of the zebrafish, because we are currently involved in a systematic study of the organization of the *Hoxb* complex in the pufferfish, which has a model compressed genome (16, 19, 20, 27).

As shown in Fig. 1*g*, construct 6 generates weak segment-restricted expression of the transgene in r3 and r5. In some embryos, the staining was localized in a narrow band in the middle of r5 or in few patches of cells of the otic vesicle (data not shown). The patterns were not as robust as those obtained with the mouse and chicken fragments, but there was highly restricted expression in r3 and r5. This result suggests that specificity has been preserved, and some aspects of the rhombomeric regulation are mediated by elements within this region. We have previously noted in mouse and chicken analysis that *Krox-20* binding sites alone are not sufficient to direct r3/r5 expression, so it is also possible that the pufferfish fragments are missing elements required for its full activity.

Conserved *Krox-20* Binding Sites in the Chicken and Pufferfish Enhancers. Together the functional results with the chicken and pufferfish genes support the idea that *Krox-20* has a conserved role in *Hox* regulation; hence, we performed an extensive set of sequence comparisons to identify conserved motifs. First, we screened for *Krox-20* binding sites in the chicken *Hoxa-2* r3/r5 regulatory region to verify our assumption that *Krox-20* might be involved in mediating the activity of the enhancer by binding directly to the DNA. Indeed, a single *Krox-20* binding site (5′-GCGTGGGTG-3′) was found in the 500-bp *Eco*RI–*Dra*I fragment, and it is identical to the high-affinity site located in the mouse *Hoxa-2* enhancer (15) (Fig. 3*B*). We have previously shown by electrophoretic mobility-shift assays that bacterially expressed Krox-20 protein will bind to this sequence from the mouse gene and form stable high-affinity complexes (15). The relative distance of these sites from the respective ATG initiation codons was similar, 2.0 kb for the mouse gene and 1.8 kb for the chicken (Fig. 3*A*). In the mouse *Hoxa-2* gene, this conserved Krox-20 binding site was shown to be necessary for enhancer activity, and the enhancer was activated by ectopic expression of *Krox-20* in transgenic mice. Therefore, by analogy to the mouse, the conserved site in the functional chicken r3/r5 enhancer is highly likely to represent a target for direct transactivation by *Krox-20*. This strongly supports a common role for *Krox-20* in vertebrates in the up-regulation of at least the *Hoxa-2* gene in r3 and r5.

In a similar manner, sequence analysis of the functional pufferfish *Hoxb-2* enhancer region revealed a single Krox-20 site (5′-GTGTGGGCG-3′) that differs by only 2 bp from the site in the mouse and chicken *Hoxa-2* genes (Fig. 3*B*). One of these bases, the T at position 2, is identical to that in the *Krox-20* binding site identified in both the human and chicken *Hoxb-2* gene. C or T at this position appear equally compatible with Krox-20 binding, as an electrophoretic mobility-shift

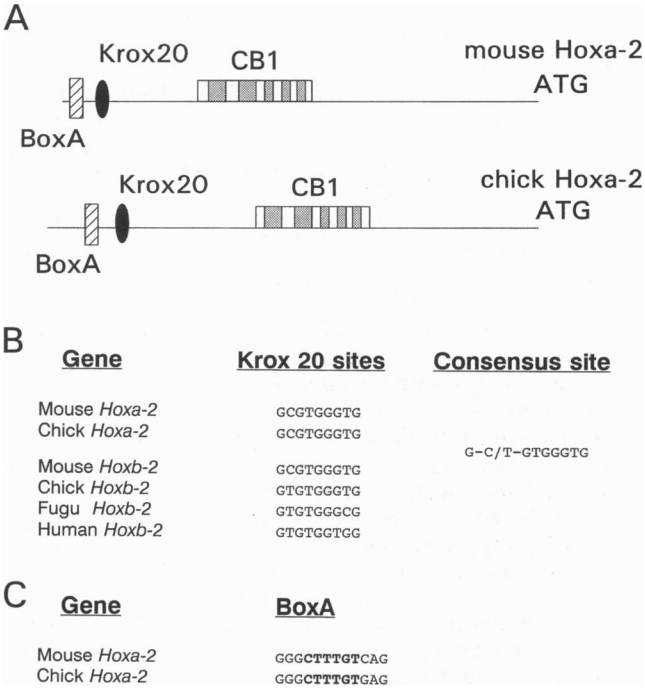

FIG. 3. Map and schematic representation of conservation between the mouse *Hoxa-2* and chicken *Hoxa-2* r3/r5 enhancer elements. (*A*) The Krox-20 binding sites and the adjacent highly conserved motif (BoxA) are indicated by a solid ellipse and a hatched rectangle, respectively. The shaded boxes in the conserved block (CB1) show the presence of five identical motifs within the stretch of homologous sequence. Note that the Krox-20 binding site and CB1 are situated at similar positions with respect to the ATGs in both genomic sequences. (*B*) Alignment of the Krox-20 binding sites from the respective genes and their consensus sequences. The human sequence (28) has TG at positions 7 and 8, which would make it a weak site. (*C*) The BoxA sequence adjacent to the Krox-20 site in the *Hoxa-2* enhancers.

assay with the mouse *Hoxb-2* sequence has shown that bacterially expressed Krox-20 can form high-affinity complexes *in vitro* (14). The remaining difference is the C at position 8, at which a T is usually present in other characterized sites; however, the human *Hoxb-2* gene is also variant in this position, containing a G. We have used double-stranded oligonucleotides corresponding to the sequence of the pufferfish *Krox-20* site as a competitor in electrophoretic mobility-shift assays with labeled versions of known high-affinity *Krox-20* sites and found that it effectively blocked complex formation (data not shown). Therefore, the pufferfish site is a valid *Krox-20* binding site.

Using the sequence information in this paper, it is possible to generate the consensus *Krox-20* site (5′-G-C/T-GT-GGGTG-3′) by comparing it to the active *in vivo* sites in our previous transgenic analysis of the *Hoxa-2* and *Hoxb-2* genes from human, mouse, chicken, and pufferfish enhancers, where available (Fig. 3*B*). During the duplication and divergence of the vertebrate *Hox* complexes, high-affinity Krox-20 binding sites have been conserved in the 5′-flanking region of the group 2 paralogous genes, again supporting the idea that there is a critical role for *Krox-20* in up-regulation of segment-restricted expression in r3 and r5 in the vertebrate hindbrain.

A Conserved Motif (BoxA) Associated with the *Krox-20* Site in *Hoxa-2*. The *Krox-20* site in the mouse *Hoxa-2* gene has been shown to be necessary for regulatory activity of the enhancer (15); however, it is not sufficient for r3/r5 expression. Recently, analysis of the mouse and chicken *Hoxb-2* gene revealed that a sequence motif (termed Box1) adjacent to the *Krox-20* binding site at a distance of 17 bp is also absolutely required for r3/r5 enhancer function in transgenic mice (29). This has opened up the possibility that other factors may participate in modulating the ability of *Krox-20* to bind to its target sites and/or cooperate with it to stimulate transactivation. Therefore, we paid particular attention to the sequences immediately surrounding the *Krox-20* site in chicken *Hoxa-2*. On the 5′ side of the *Krox-20* sites, 3 bp upstream in the mouse gene and 8 bp upstream in the chicken, there is a region with 11 of the 12 bp identical. We have referred to this element as BoxA (Fig. 3 *A* and *C*) and looked for motifs related to those observed in the *Hoxb-2* Box1 motif. The precise sequence of the Box1 element in *Hoxb-2* is not present; however, within BoxA we noted that there is smaller core of 6 bp (5′-CTTTNN-3′ with a preference of GT for the last 2 bp) that is partially conserved with *Hoxb-2* Box1 sites, although the spacing of this core is different. Searches using this site have revealed no homology with binding sites of known transcription factors. Functional analysis will be required to determine if these BoxA motifs are required for *Krox-20* activity of the *Hoxa-2* enhancers.

Other Conserved Blocks in the *Hoxa-2* r3/r5 Enhancers. In previous experiments, which demonstrated that *Krox-20* is a direct regulator of mouse *Hoxb-2* expression in r3 and r5, it was shown that a 578-bp fragment containing the high-affinity *Krox-20* site and the adjacent Box1 motif is not sufficient in single copy to mediate r3/r5 expression of a *lacZ* reporter gene (14, 29). However, addition of a 122-bp 5′-flanking region linked to this 578-bp fragment was able to direct high levels of r3/r5 transgene expression. Furthermore, multimerization of the 578-bp region alone also generated the proper r3/r5 staining (14). These experiments suggested that the specificity for restricted expression in r3 and r5 probably resides in the 578-bp region containing the *Krox-20* sites, but that other elements in the 122-bp region may be involved cooperatively or independently in regulating the efficiency of these *Krox-20*-mediated activities. Therefore, we were interested in searching for additional conserved blocks in larger regions surrounding the *Krox-20* binding sites in the mouse and chicken *Hoxa-2* genes.

Surprisingly, the overall sequence similarity of a 1-kb region containing the *Krox-20* sites was relatively low between the chicken and mouse *Hoxa-2* enhancers. There was one conserved block (CB1) of about 200 bp downstream of the *Krox-20* sites with 75% identity between mouse and chicken (Fig. 3*A*). The identity within the 200 bp was concentrated in five colinear subdomains of 33, 29, 12, 32, and 22 bp. In the mouse, the CB1 domain lies in the minimal r3/r5 enhancer, which generates high levels of restricted expression in both r3 and r5. However, in the chicken gene, CB1 lies in the larger 1.2-kb r3/r5 enhancer but not within the minimal 500-bp *Eco*RI–*Dra*I enhancer. Since both have weak r3 expression, CB1 cannot be used to account for the lower levels of staining in r3 of the chicken enhancer. However, as noted above, the 500-bp *Eco*RI–*Dra*I enhancer (construct 4) was less efficient in transgenic assays when compared with the full 1.2-kb enhancer (construct 2). Therefore, it is tempting to speculate that CB1 might be involved in modulating the efficiency of the r3/r5 enhancer, and it will be important to specifically test the functional role of motifs in this domain. We also looked for CB1-related sequences in the 1.8-kb pufferfish 5′ *Hoxb-2* enhancer but found no matches. If such homologous regions exist, they are outside of the fragment we tested, which might account for the relatively low efficiency of transgene expression with the pufferfish enhancer.

DISCUSSION

The experiments presented in this paper serve to reinforce the value of evolutionary approaches in studying the developmental regulation of vertebrate *Hox* genes. Regulatory regions from the chicken and pufferfish *Hox* genes are able to interact with upstream factors, such as *Krox-20*, in the heterologous

mouse system and generate the appropriate patterns of segmentally restricted expression of a *lacZ* reporter gene. Hence, in addition to the extensive conservation of *Hox* proteins and a similar chromosomal organization of the *Hox* clusters, there appears to be a considerable degree of conservation in the underlying molecular mechanisms that regulate their expression. Although we have focused on the conserved aspects of segmental regulation to find common components, it will be equally important to investigate those patterns that differ between species, as this could represent a means for the generation of evolutionary diversity.

***Krox-20* Has a Common Role.** In the mouse, *Krox-20* is clearly an important component in regulating the events that lead to the full morphogenetic development of r3 and r5. Both *Hoxb-2* and *Hoxa-2* are direct targets of *Krox-20* in r3 and r5, and the loss of their function in this domain could be responsible for the abnormalities observed in the *Krox-20* mutants (12, 13). Loss of *Hoxa-2* alone in targeted mouse mutants has no detectable segmental abnormalities (30, 31), although this could be a consequence of functional compensation by *Hoxb-2*, and it will be essential to make *Hoxa-2/Hoxb-2* double mutants to determine if these are the sole components. It is important to note that a number of other genes, such as the *Sek1* segmentally expressed tyrosine kinase of the EPH family, are also expressed at high levels in r3 and r5 at the same time as *Hoxa-2* and *Hoxb-2* (32, 33). Therefore, *Krox-20* could play a more general role in regulating genes displaying r3/r5-restricted expression.

Results from these experiments reinforce the concept that *Krox-20* plays a similar role in patterning the hindbrain of other vertebrates. *Krox-20* expression in r3 and r5 is conserved in chick, fish, frog, and mouse embryos (26, 34–36), although the dynamics of appearance and down-regulation varies between species. Furthermore, the boundaries and patterns of *Hox* gene expression in these species are identical to the mouse, to the extent currently tested. Our regulatory analysis demonstrates that the similar patterns of *Hox* expression are generated by conserved r3/r5 regulatory enhancers in the chicken and pufferfish genes. The sequence comparisons identified conserved high-affinity *Krox-20* binding sites that would provide targets for the transactivation by *Krox-20*. The presence of nearly identical Krox-20 binding sites in six of the genes we have examined shows that the major mechanisms generating r3/r5-restricted expression are the same between vertebrates.

Although our experiments highlight the conserved role of *Krox-20* in regulating early aspects of the patterning of regional identity, we do not favor the idea that *Krox-20* is a segmentation gene itself. The phenotypes in the *Krox-20* mutant embryos show that r3 and r5 form properly but fail to be maintained (12, 13). This might be achieved through its regulation of the *Hox* genes, because they could have a role in proliferative related processes, in a manner similar to that suggested in limb development (37). However, there may be a number of other *Krox-20* targets that are involved in controlling the maintenance of the rhombomeres independent of the *Hox* segment identity cascade. Because the expression patterns of *Krox-20* itself display segment-restricted domains mapping to future r3 and r5 in other vertebrates (26, 34–36), we would suggest that it is the genes implicated in establishing these earlier patterns that are the true segmentation genes in vertebrates.

Differences in r3 and r5 Regulation. We have stressed the similar patterns of transgene expression mediated by the chicken *Hoxa-2* and pufferfish *Hoxb-2* enhancers. With respect to regulation in r5, in chicken we saw strong and uniform staining of the reporter gene throughout all the cells of this segment, and the timing of appearance corresponded to that of the endogenous mouse *Hox* gene and *Krox-20*. In contrast, the staining in r3 was patchy, weak, and temporally delayed. In a recent transgenic analysis of a chicken *Hoxb-2* r3/r5 regulatory region, we also noted that *lacZ* reporter expression was weak in r3 and strong in r5 (29). This result illustrates that there is a consistent difference in the ability of the chicken *Hoxa-2* and *Hoxb-2* enhancers to effectively function in r3 as compared with r5.

This difference cannot be attributed to *Krox-20* alone, because we know that it is present at the proper times and can interact with the site for regulation in r5. These high-affinity Krox-20 binding sites in the chicken genes are also identical to those in the mouse enhancers, and transgenic analysis of the mouse regulatory regions indicates that the *Krox-20* sites are necessary but not sufficient for r3/r5 activity. Therefore, additional factors must be involved in mediating the ability of *Krox-20* to activate r3 expression, and the cofactor requirements for r3 and r5 appear to be different. Our preliminary mutational analysis of the mouse *Hoxa-2* enhancer also suggests that separate cis-acting elements in addition to the *Krox-20* binding site may be required for regulation in different rhombomeres. Hence, even though *Krox-20* shows restricted expression in r3 and r5 and is necessary for regulating the *Hox* target genes, the presence of additional components restricts its activity and provides another level of control in patterning the regional identity of hindbrain segments.

Search for *Krox-20*-Associated Motifs. Our comparative sequence analysis was directed toward identifying conserved blocks of sequence between species that might correspond to cis-acting sites for these additional factors. However, the *Hoxa-2* analysis was unlike the comparative studies of *Hoxb-4* and *Hoxb-1*, where we found a few large (100–200 bp) domains of sequence identity outside the coding regions that corresponded to functionally conserved regulatory regions (16–20). With the exception of CB1, there were no large blocks of sequence identity in the 2-kb enhancers. In fact, we had to rely on our developing knowledge of the *Krox-20* sites defined by previous *in vivo* analysis to identify the Krox-20 binding sites themselves. This suggested that the additional factors must be interacting with small individual elements dispersed in the region rather than with a set of tightly clustered sites.

Adjacent to the Krox-20 site in the chicken *Hoxa-2* enhancer, 8 bp upstream we found a 12-bp motif in which 11 of the 12 bases were conserved in the mouse enhancer. This sequence was called BoxA and was similarly positioned 3 bp upstream in the mouse gene (Fig. 3). It is tempting to assume that this conserved sequence might be functionally relevant because of its proximity to the Krox-20 binding sites. In support of this idea, we have recently demonstrated that there is a short conserved motif (Box1) in a region flanking the Krox-20 binding sites in the chicken and mouse *Hoxb-2* r3/r5 enhancers (29). This element is spaced 17 bp upstream and mutational analysis has indicated that it is required for enhancer activity in transgenic mice. The Box1 motif does not match a consensus sequence for any known transcription factor binding site described to date, nor does the sequence precisely correspond to the BoxA motif we found in this study. However, we noted that there is a core of similarity, 5′-CTTTNN-3′ with a preference of GT for the last two bases, that is present in both BoxA and Box1. On this basis, BoxA is a reasonable candidate for one of the cis-acting elements involved in cooperation with *Krox-20*, and it will be essential to directly test its functional role in a manner similar to that used to test Box1.

The Relative Degree of Sequence Conservation Within Regulatory Regions. The four vertebrate *Hox* complexes have arisen by duplication and divergence from a common ancestral cluster (reviewed in refs. 7, 9, 10, 37, and 38). In the cases like those we have detailed above, in which paralogous genes are expressed in similar domains and regulated by conserved enhancers, we expected that there would be considerable sequence similarity between the 5′-flanking regions of the chicken, mouse, and human genes. For example, how confident could we be of the putative role of the BoxA motif if there were

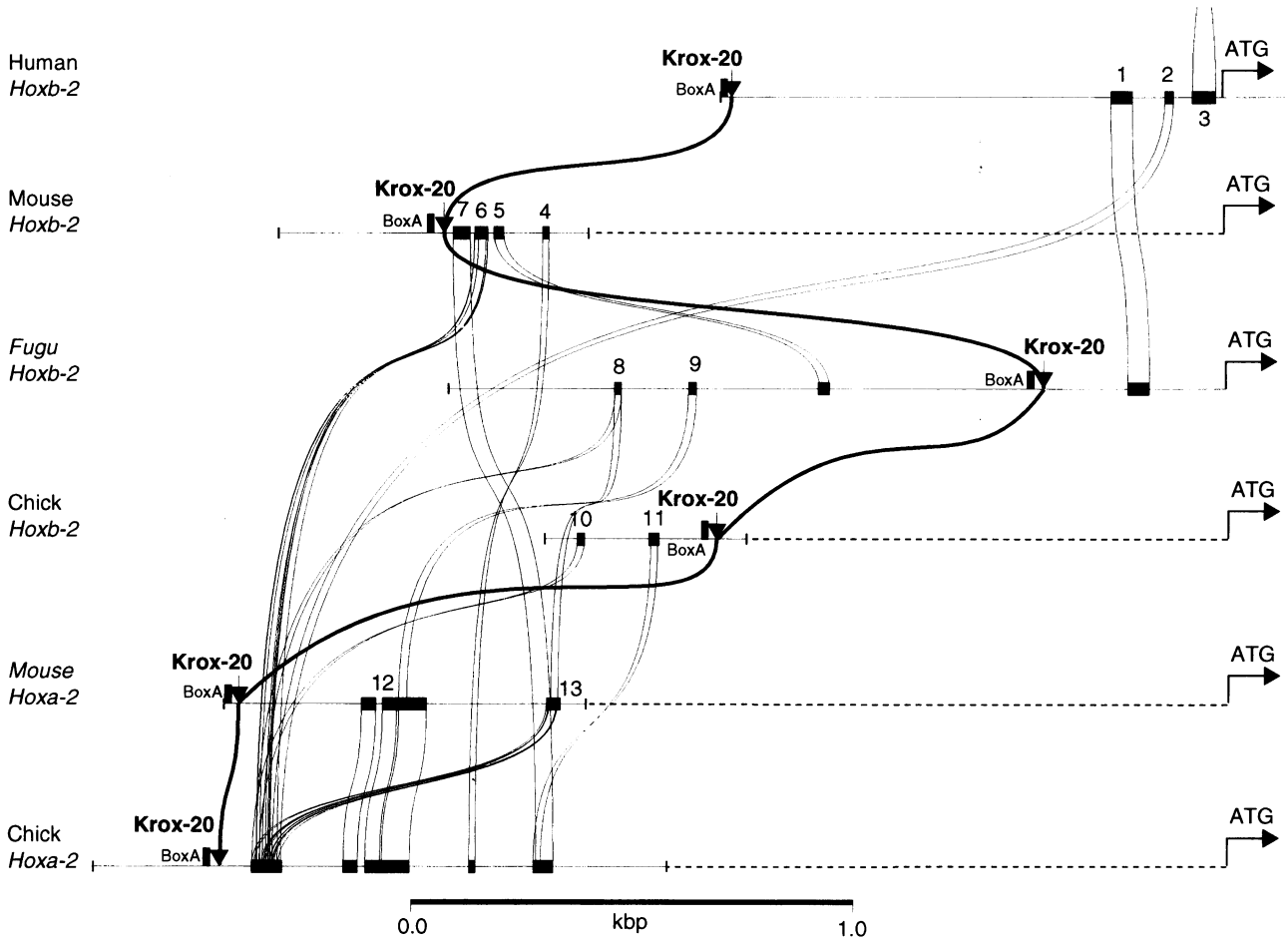

FIG. 4. Schematic alignment of the r3/r5 enhancer elements of group 2 Hox genes from mouse, chicken, pufferfish, and human. A multiple graph comparing the r3/r5 regulatory elements, where curves connect homologous blocks (solid squares) of sequence in the regulatory domains. The Krox-20 binding sites (indicated by a solid triangle) and the BoxAs (indicated by a solid rectangle) are situated at similar locations with respect to the ATGs in all six fragments and joined with a thicker curve. Block 12 is the homology CB1 in the mouse and chicken Hoxa-2 genes described in Fig. 3. The solid boxes (blocks 1–13) represent short stretches of at least 70% sequence identity dispersed along the enhancer elements of all four species. Regions of the mouse and chicken genomic DNAs between the ATGs and the enhancers that have not been sequenced are depicted with a broken line. For ease of interpretation, the Box1 motif in the *Hoxb-2* genes that is related to BoxA in *Hoxa-2* is noted as BoxA.

a large number of 12-bp runs of sequence identity throughout the flanking regions. Therefore, we felt it might be necessary to compare regions from even more evolutionary distant vertebrates, such as the pufferfish, to minimize nonfunctional sequence similarities and more readily detect functional cis-acting regulatory elements.

Fig. 4 illustrates the alignment of the sequenced enhancer regions from the *Hoxb-2* and *Hoxa-2* genes of human, mouse, chicken, and pufferfish. The relative positions of the high-affinity Krox-20 binding sites are indicated (Krox-20) along with the associated BoxA motif. For ease in interpretation and because of the related core sequence, we have denoted the Box1 motif of the *Hoxb-2* genes as BoxA. With the exception of block 12, representing the region CB1 described earlier (Fig. 3*A*), the blocks of sequence conservation (blocks 1–13) reflect 70% similarity between a number of short stretches of sequence dispersed in the enhancers. Many of the conserved blocks are not organized in a colinear manner, as their positions can vary between different enhancers. For example, blocks 8 and 9 are upstream of the Krox-20 site in the pufferfish enhancer and downstream of the site in the chicken and mouse enhancers. The relative distances of several blocks of homologies also vary considerably with respect to their proximity to the Krox-20 binding site (for example, see refs. 2, 7, and 13). These alignments reveal that there is considerably less conservation in the flanking regions of these *Hox* genes than anticipated. Therefore, we are not encumbered with a large number of blocks of homology, many of which are likely to be nonfunctional sequences. If important functional domains had resided within large conserved blocks of this region, they would have been easily identified, as we found in the case of *Hoxb-4* (20). However, with the exception of CB1, the relevant motifs appear small and/or divergent and their position can vary between species. This makes the identification of functional domains by evolutionary sequence comparisons very difficult.

Despite these limitations, our evolutionary comparisons of the group 2 paralogs have been extremely useful in showing that *Krox-20* plays a conserved role in hindbrain segmentation. We are beginning to identify additional cis-acting sequences that cooperate in these regulatory processes, and it will be important to isolate the factors that bind to them to determine if they have restricted or more general/ubiquitous roles. Clearly, the similarity in hindbrain morphology and organization between vertebrates will greatly aid in building a complete picture of the regulatory cascade, because it is underscored by a conserved set of molecular mechanisms.

We thank Joe Brock for computer photography, and Dr. Mai-Har Sham and members of the Krumlauf laboratory for valuable discussion. S.N. was supported by a European Commission Biotechnology grant. M. Maconochie and S.A. were supported by Medical Research Council Training Fellowships, and M. Manzanares was supported by

a Human Frontiers in Science Program Fellowship. This work was partially supported by Human Frontiers in Science Program and European Economic Community Biotechnology Network grants to R.K. and grants from Institut National de la Santé et de la Recherche Médicale, European Economic Community, Association pour la Recherche Contre le Cancer, Ligue Nationale Française Contre le Cancer, Association Française Contre les Myopathies, and Association pour la Recherche sur la Sclrose en Plaque to P.C.

1. Le Douarin, N. (1983) *The Neural Crest* (Cambridge Univ. Press, Cambridge, U.K.).
2. Wilkinson, D. G. (1993) *BioEssays* **15,** 499–505.
3. Fraser, S., Keynes, R. & Lumsden, A. (1990) *Nature (London)* **344,** 431–435.
4. Lumsden, A. & Keynes, R. (1989) *Nature (London)* **337,** 424–428.
5. Noden, D. (1983) *Dev. Biol.* **96,** 144–165.
6. Noden, D. (1988) *Development (Cambridge, U.K.)* **103,** Suppl., 121–140.
7. Krumlauf, R. (1992) *BioEssays* **14,** 245–252.
8. Krumlauf, R. (1993) *Trends Genet.* **9,** 106–112.
9. McGinnis, W. & Krumlauf, R. (1992) *Cell* **68,** 283–302.
10. Krumlauf, R. (1994) *Cell* **78,** 191–201.
11. Keynes, R. & Krumlauf, R. (1994) *Annu. Rev. Neurosci.* **17,** 109–132.
12. Schneider-Maunoury, S., Topilko, P., Seitanidou, T., Levi, G., Cohen-Tannoudji, M., Pournin, S., Babinet, C. & Charnay, P. (1993) *Cell* **75,** 1199–1214.
13. Swiatek, P. J. & Gridley, T. (1993) *Genes Dev.* **7,** 2071–2084.
14. Sham, M. H., Vesque, C., Nonchev, S., Marshall, H., Frain, M., Das Gupta, R., Whiting, J., Wilkinson, D., Charnay, P. & Krumlauf, R. (1993) *Cell* **72,** 183–196.
15. Nonchev, S., Vesque, C., Maconochie, M., Seitanidou, T., Ariza-McNaughton, L., Frain, M., Marshall, H., Sham, M. H., Krumlauf, R. & Charnay, P. (1996) *Development (Cambridge, U.K.)* **122,** 543–554.
16. Popperl, H., Bienz, M., Studer, M., Chan, S. K., Aparicio, S., Brenner, S., Mann, R. S. & Krumlauf, R. (1995) *Cell* **81,** 1031–1042.
17. Studer, M., Popperl, H., Marshall. H., Kuroiwa, A. & Krumlauf, R. (1994) *Science* **265,** 1728–1732.
18. Marshall, H., Studer, M., Popperl, H., Aparicio, S., Kuroiwa, A., Brenner, S. & Krumlauf, R. (1994) *Nature (London)* **370,** 567–571.
19. Aparicio, S., Morrison, A., Gould, A., Gilthorpe, J., Chaudhuri, C., Rigby, P. W. J., Krumlauf, R. & Brenner, S. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 1684–1688.
20. Morrison, A., Chaudhuri, C., Ariza-McNaughton, L., Muchamore, I., Kuroiwa, A. & Krumlauf, R. (1995) *Mech. Dev.* **53,** 47–59.
21. Wilkinson, D. & Green, J. (1990) in *The Practical Approach Series: Postimplantation Mouse Embryos: A Practical Approach*, eds. Copp, A. J. & Cockcroft, D. L. (IRL, Oxford), pp. 155–171.
22. Whiting, J., Marshall, H., Cook, M., Krumlauf, R., Rigby, P. W. J., Stott, D. & Allemann, R. K. (1991) *Genes Dev.* **5,** 2048–2059.
23. Baxendale, S., Abdulla, S., Elgar, G., Buck, D., Berks, M., Micklem, G., Durbin, R., Bates, G., Brenner, S., Beck, S. & Lehrach, H. (1995) *Nat. Genet.* **10,** 67–76.
24. Wilkinson, D. G., Bhatt, S., Cook, M., Boncinelli, E. & Krumlauf, R. (1989) *Nature (London)* **341,** 405–409.
25. Prince, V. & Lumsden, A. (1994) *Development (Cambridge, U.K.)* **120,** 911–923.
26. Oxtoby, E. & Jowett, T. (1993) *Nucleic Acids Res.* **21,** 1087-1095.
27. Brenner, S., Elgar, G., Sandford, R., Macrae, A., Venkatesh, B. & Aparicio, S. (1993) *Nature (London)* **366,** 265–268.
28. Vieille-Grosjean, I. & Huber, P. (1995) *J. Biol. Chem.* **270,** 4544–4550.
29. Vesque, C., Maconochie, M., Nonchev, S., Ariza-McNaughton, L., Kuroiwa, A., Charnay, P. & Krumlauf, R. (1996) *EMBO J.*, in press.
30. Gendron-Maguire, M., Mallo, M., Zhang, M. & Gridley, T. (1993) *Cell* **75,** 1317–1331.
31. Rijli, F. M., Mark, M., Lakkaraju, S., Dierich, A., Dolle, P. & Chambon, P. (1993) *Cell* **75,** 1333–1349.
32. Becker, N., Seitanidou, T., Murphy, P., Mattei, M.-G., Topilko, P., Nieto, M. A., Wilkinson, D., Charnay, P. & Gilardi-Hebenstreit, P. (1994) *Mech. Dev.* **47,** 3–18.
33. Nieto, M. A., Gilardi-Hebenstreit, P., Charnay, P. & Wilkinson, D. (1992) *Development (Cambridge, U.K.)* **116,** 1137–1150.
34. Bradley, L. C., Snape, A., Bhatt, S. & Wilkinson, D. G. (1992) *Mech. Dev.* **40,** 73–84.
35. Nieto, M. A., Bradley, L. C. & Wilkinson, D. G. (1991) *Development (Cambridge, U.K.)* **2,** Suppl., 59–62.
36. Wilkinson, D. G., Bhatt, S., Chavrier, P., Bravo, R. & Charnay, P. (1989) *Nature (London)* **337,** 461–464.
37. Duboule, D. & Morata, G. (1994) *Trends Genet.* **10,** 358–364.
38. Boncinelli, E., Simeone, A., Acampora, D. & Mavilio, F. (1991) *Trends Genet.* **7,** 329–334.