Conserved expression domains for genes upstream and within the *HoxA* and *HoxD* clusters suggests a long-range enhancer existed before cluster duplication

Jessica A. Lehoczky,[a] Melissa E. Williams,[a] and Jeffrey W. Innis[a,b,*]

[a]Department of Human Genetics and
[b]Department of Pediatrics, Division of Genetics, University of Michigan, Ann Arbor, MI 48109, USA

*Author for correspondence (email: innis@umich.edu)

SUMMARY   The posterior *HoxA* and *HoxD* genes are essential in appendicular development. Studies have demonstrated that a "distal limb enhancer," remotely located upstream of the *HoxD* complex, is required to drive embryonic autopod expression of the posterior *Hox* genes as well as the two additional non-*Hox* genes in the region: *Evx2* and *Lnp*. Our work demonstrates a similar mode of regulation for *Hoxa13* and four upstream genes: *Evx1*, *Hibadh*, *Tax1bp*, and *Jaz1*. These genes all show embryonic (E11.5–E13.5) distal limb and genital bud expression, suggesting the existence of a nearby enhancer influencing the expression of a domain of genes. Comparative sequence analysis between homologous human and mouse genomic sequence upstream of *Hoxa13* revealed a remote 2.25-kb conserved noncoding sequence (mmA13CNS) within the fourth intron of the *Hibadh* gene.

mmA13CNS shares a common 131-bp core identity within a conserved noncoding sequence upstream of *Hoxd13*, which is located within the previously identified distal limb enhancer critical region. To test the function of this conserved sequence, we created mmA13CNS-*Hsp86-lacZ* transgenic mice. mmA13CNS directed a wide range of tissue expression, including the central nervous system, developing olfactory tissue, limb, and genital bud. Limb and genital bud expression directed by mmA13CNS is not identical to the patterns exhibited by *Hoxa13/Evx1/Hibadh/Tax1bp1/Jaz1*, suggesting that mmA13CNS is not sufficient to fully recapitulate their expression in those tissues. The *Evx1*- and *Evx2*-like central nervous system expression observed in these mice suggests that the long-range regulatory element(s) for the *Hox* cluster existed before the cluster duplication.

INTRODUCTION

Hox genes encode essential transcription factors for embryonic development of the axial and appendicular body plan of metazoans (McGinnis and Krumlauf 1992; Veraksa et al. 2000). The *Hox* genes are located in genomic clusters with genes 3′ in the cluster expressed earliest in development and most anterior in the developing axial body plan and genes 5′ in the cluster expressed later in development and more posterior in the organism (Duboule and Morata 1994). *Branchiostoma floridae* (amphioxus), an extant cephalochordate related to the pre-*Hox* duplication vertebrate ancestor, has one *Hox* cluster comprised of 14 genes (Ferrier et al. 2000), yet over vertebrate evolution the genomic clusters duplicated such that mammals have four *Hox* clusters, *HoxA* through *HoxD*, with a total of 39 genes spread among 13 paralogous groups (Ruddle et al. 1994).

Among numerous sites of expression, mammalian development of the appendicular skeleton and external genitalia relies on the 5′ *Hox* genes of paralog groups A and D (Fro-

mental-Ramain et al. 1996; Mortlock et al. 1996; Kondo et al. 1997; Warot et al. 1997). In the *HoxD* cluster, genes *Hoxd10* through *Hoxd13* have been shown to be necessary for proper limb development in mice (Zakany et al. 1997). In the *HoxA* cluster, *Hoxa13* is essential in distal limb development, as shown by genetically engineered *Hoxa13* null alleles (Fromental-Ramain et al. 1996; Stadler et al. 2001), the mouse mutant Hypodactyly (Mortlock et al. 1996), and patients with hand–foot–genital syndrome (Mortlock and Innis 1997; Goodman et al. 2000) (OMIM #140000). During mouse development, *Hoxd10* through *Hoxd13* are expressed in the mesenchyme of the distal fore- and hindlimb autopods as well as the genital tubercle from mouse embryonic days 10.5 through 13.5 (E10.5–E13.5). This expression is comparable with that of the *Hoxd13* paralog, *Hoxa13*, which shows similar distally restricted limb bud and genital bud expression during embryogenesis.

Investigation into the expression domains of genes 5′ to *Hoxd13* or *Hoxa13* revealed that additional non-*Hox* genes share their distal limb and genital bud expression patterns

during embryonic development in mice. *Evx2* and *Lunapark* (*Lnp*) are located upstream of *Hoxd13*, and both have mesenchymal distal limb and genital bud expression as part of their expression domains (Spitz et al. 2003). *Evx1* is located 5′ of *Hoxa13* and similarly has distal and genital bud expression during embryonic development (Bastian and Gruss 1990; this study). In addition, *Evx1* and *Evx2* share additional embryonic expression domains in the developing central nervous system where neither 5′ *HoxA* nor 5′ *HoxD* genes are expressed. The conserved distal limb and genital bud expression domains of the posterior *HoxD* and 5′ genes are regulated by a global element(s) located further 5′ of *Lnp* (Spitz et al. 2001).

Given that the Hox complexes were derived by duplication of an ancestral group of genes, we hypothesized that a regulatory mechanism similar to that proposed for posterior *HoxD* genes, which minimally involves enhancers acting over a domain of genes to direct expression in the distal limb and the genital bud, also exists for *HoxA*. Herein, we examine the spatial and temporal expression pattern of six genes upstream of *Hoxa13*: *Evx1*, *Hibadh*, *Tax1bp1*, *Jaz1*, *Crebpa*, and *KIAA0644*. We also use upstream *HoxA* and *HoxD* sequence comparisons to identify a highly conserved noncoding sequence (CNS) upstream of *Hoxa13* and test the function of this element in transgenic *lacZ* reporter mice.

MATERIALS AND METHODS

Sequence alignments and analysis

Human and mouse genomic regions 5′ of *HOXA13* and *HOXD13* were obtained from the Ensembl Genome Browser (www.ensembl.org). Human DNA sequence was obtained from Build 34, and mouse DNA sequence was obtained from Build 30. Genomic sequences were Repeatmasked (Smit and Green, unpublished results; http://www.repeatmasker.org) and subjected to multiple species alignments using advanced Pipmaker and Multipipmaker with default settings (Schwartz et al. 2000; http://bio.cse.psu.edu/pipmaker). Conserved noncoding sequences (CNSs) are referred to frequently in this study and where discussed, "central nervous system" is not abbreviated. Sequences homologous to mmA13CNS were obtained by BLAST comparison (Altschul et al. 1990) against NCBI GenBank databases nr, htg, gss, and wgs (www.ncbi.nlm.nih.gov/Genbank).

Transgenic mice and *LacZ* staining

The original *Hsp68-lacZ* vector (Kothary et al. 1989; DiLeone et al. 1998) was modified to include a *Not*I-*Sfi*I-*Not*I cloning site. We obtained this modified vector from Dr. Doug Mortlock (Vanderbilt University, Nashville, TN, USA). mmA13CNS was amplified by polymerase chain reaction from C57Bl/6 mouse genomic DNA with primers 5′-ATGTGCTCTGCCACTCTCTG and 5′-TGGAATTGGCCTTACGAAAA. The integrity of this element was confirmed by DNA sequencing and subsequently cloned into the *Not*I site of the *Hsp68-lacZ* vector. The transgene was linearized, and the vector backbone was removed. Transgene DNA was injected into pronuclei of fertilized B6D2 F2 eggs and implanted into pseudopregnant CD1 females. Embryos were harvested for β-galactosidase staining at E12.5, and amniotic membranes were collected for transgene genotyping. *LacZ* staining was performed by standard protocol (Hogan et al. 1994). Embryos used for sections were dehydrated in ethanol/phosphate-buffered saline and subsequently embedded in paraffin by standard protocol. Embryos were sectioned at 6 or 10 μm, as noted.

In situ hybridization

Hibadh, *Tax1bp1*, *Crebpa*, and *Jaz1* cDNAs were amplified from an E12.5 C57Bl/6 mouse limb bud cDNA library and cloned into polymerase chain reaction-4 Topo vector (Invitrogen, Carlsbad, CA, USA). Two clones/probes were made for *Hibadh* and pooled for in situ hybridization. *Hibadh* primer pairs were as follows: hibadhex7.F 5′-CAGGCCTTGACCCAAAACTA and hibadh_3′UTR.R 5′-TCATCATTGGCTCTGAAACAG, mm_hibadh.F 5′-GGATTCATTGGACTGGGAAA and mm_hibadh.R 5′-ATGGGTGTGCTGGCTAAAC. All other primer pairs were as follows: jaz1.F 5′-GGAGTCAGACAGTGATGAGTCC and jaz1.R 5′-GGTTAAACAATATGCAACATGCC, mmCrebpa.F 5′-TTATCTCAGACCTGGAAGACG and mmCrebpa.R 5′-TTTCCCTAGTGTCCCCACAG, Tax1bp1ex15.F 5′-GCCCTCAACAAGTCTCAAGG and Tax1bp1_3′UTR.R 5′-CAATCGCCAAGGTGTACAGA. The *Evx1* clone was kindly provided by Dr. Gail Martin (UCSF, San Francisco, CA, USA). Digoxygenin-labeled anti-sense RNA probes were generated as previously described (Post and Innis 1999). In situ hybridization was carried out as previously described (Bober et al. 1994) except embryos were developed using BM purple substrate (Roche Applied Science, Indianapolis, IN, USA).

RESULTS

We hypothesized that the mechanism regulating distal limb bud and genital bud expression in posterior *HoxD* and 5′ genes may have been conserved on the *HoxA* chromosome. We sought to characterize the embryonic RNA expression patterns of the genes within 1.5 megabases upstream of *Hoxa13*. Six genes were identified within this interval in human and mouse: *Evx1*, *Hibadh*, *Tax1bp1*, *Jaz1*, *Crebpa*, and *KIAA0644* (Fig. 1). Notably, the *HoxA* cluster through *Crebpa* is a conserved syntenic block of genes from at least bony fish through humans, without apparent gene insertions or deletions. It would be of interest to examine gene context upstream of the *HoxA* clusters in lampreys and sharks; however, this genomic sequence is not available. *Evx1*, the first gene 5′ of the *HoxA* cluster, was originally identified to be a homolog of *Drosophila even-skipped* (Bastian and Gruss 1990). *Evx1* expression domains are closely correlated to that of *Evx2*, the first gene upstream of *Hoxd13*, with expression in the developing neural tube, hindbrain, distal limbs, and genital bud (Bastian and Gruss 1990; Dollé et al. 1994) (Fig. 1, E–H). The expression seen for *Evx1* in the distal limbs and
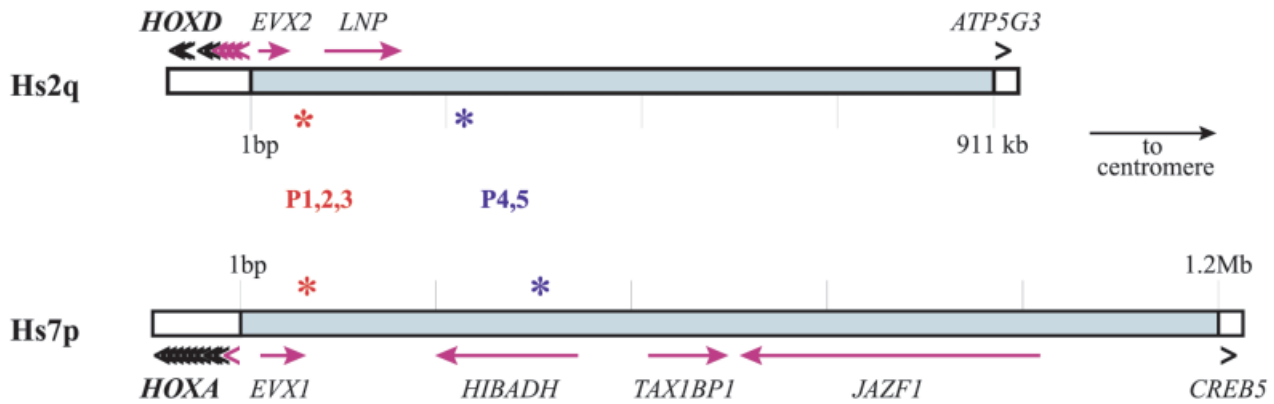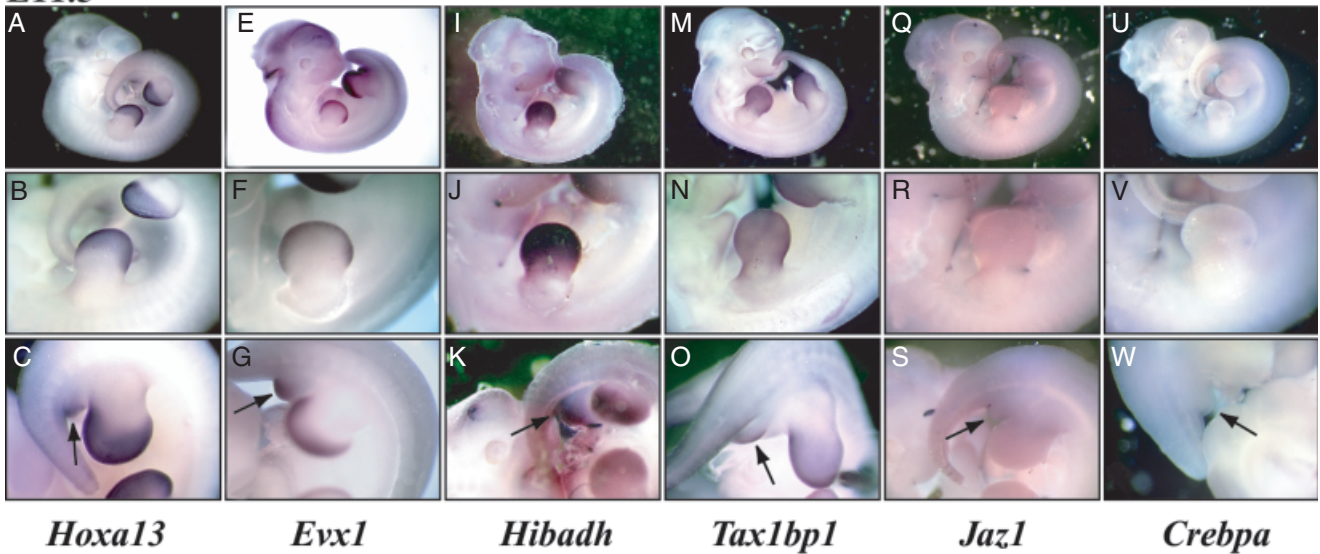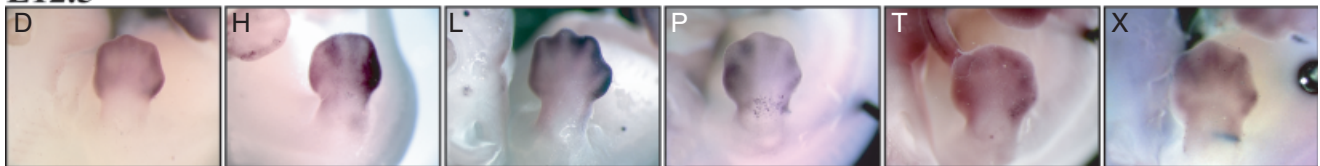
Fig. 1. Genomic organization and expression domains of genes upstream of *HoxA*. Top shows 911 kb of human genomic sequence 5′ of *HOXD13*, where base 1 is the first base before the start of *HOXD13* transcription and base 911,179 is the last base before the start of *ATP5G3* transcription. Below this diagram 1.2 Mb of human genomic sequence 5′ of *HOXA13* is drawn to scale. Base 1 is the first base before the start of *HOXA13* transcription and base 1.2 Mb is the first base before the start of *CREB5*. Arrows below gene names denote the direction of transcription and those in pink share common embryonic distal limb expression patterns. Shaded gray portions represent sequence used for comparative analysis. Red asterisks denote the relative location of paraCNSs 1, 2, and 3. Blue asterisks show the relative location of paraCNS4/5 and A13CNS/D13CNS, respectively. Bottom shows whole-mount RNA expression domains of *Hoxa13* (A–D), *Evx1* (E–H), *Hibadh* (I–J), *Tax1bp1* (M–P), *Jaz1* (Q–T), and *Crebpa* (U–X) for mouse embryos. A, E, I, M, Q, and U show E11.5 right sides, and B, F, J, N, R, and V show close-up of right E11.5 forelimbs. Arrows in C, G, K, O, S, and W show E11.5 genital bud expression. D, I, P, T, and X show E12.5 forelimb expression, and H shows E12.5 hindlimb, as *Evx1* begins turning off at E12.5 in the forelimb.

genital bud are similar to that of *Hoxa13* in terms of spatial and temporal characterization (Fig. 1, A–D).

   3-Hydroxyisobutyrate dehydrogenase (*HIBADH*) is located 325 kb 5′ of human *HOXA13*. *HIBADH* is expressed prima-

rily in the adult liver and is involved in valine catabolism (Rougraff et al. 1988). There is no previous evidence of embryonic expression for this gene. *Hibadh* anti-sense RNA in situ hybridization for mouse E10.5 through E13.5 revealed a

mesenchymal staining pattern highly similar to that of *Hoxa13*, with mesenchymal distal limb and genital bud staining (Fig. 1, I–J). *Tax1 binding protein 1* (*TAX1BP1*), located 540 kb 5′ of *HOXA13*, has been shown in vitro to be a substrate for caspase-3-like proteases in the tumor necrosis factor–induced apoptotic pathway (De Valck et al. 1999). Although no embryonic expression profile has been reported for this gene, RNA in situ hybridization shows a distal limb and genital bud staining pattern similar to that of *Hoxa13* (Fig. 1, M–P). *Juxtaposed with another zinc-finger gene 1* (*Jaz1*; human, *JAZF1*), located 630 kb 5′ of human *HOXA13*, has only been reported in the context of JAZF1/JJAZ fusion protein expression in endometrial stromal neoplasms (Koontz et al. 2001). Evaluation of the developmental RNA expression profile of *Jaz1* for mouse E10.5 and 11.5 shows low-level, ubiquitous, embryonic expression including the limb and the genital bud (Fig. 1, Q–S). This ubiquitous expression is reproducible and not correlated with time in developing reagent. At E12.5 and E13.5, the expression pattern is dramatically restricted to the interdigital mesenchyme (Fig. 1T), central nervous system, and genital bud (data not shown).

Two additional genes 5′ of *Hoxa13* were evaluated for their relative expression during E10.5 through E13.5: *Crebpa* and *KIAA0644*. *cAMP responsive element binding protein* (*Crebpa*; human, *CREB5*) shows ventral neural tube, forebrain, and faint autopod staining dissimilar to *Hoxa13* and no genital bud expression at E11.5 (Fig. 1, U–W). *Crebpa* displays an interdigital staining tightly confined to the edges of the condensing mesenchyme at stages E12.5 (Fig. 1X) and E13.5, in addition to forebrain staining (data not shown). *KIAA0644*, originally isolated from a human brain cDNA library (GenBank accession number AB014544), does not show the characteristic *Hoxa13* expression pattern. KIAA0644 expression was observed in the proximal, not distal, limb as well as the forebrain (data not shown).

To identify candidate regulatory elements necessary for the distal limb and genital bud expression pattern observed for *Hoxa13* and the four upstream genes, comparative sequence analysis was performed using human and mouse genomic sequence. Comparison of approximately 900 kb of human genomic sequence 5′ of *HOXD13* with 1.2 Mb of human sequence 5′ of *HOXA13* yielded five conserved noncoding

sequences, and identical results were obtained by comparison with the paralogous mouse regions (Table 1). All five CNSs were core sequence identities, which fell within larger homologous CNSs found between the human and mouse homologous cluster comparison. The five elements were further characterized by their locations. We define a paraCNS as conserved noncoding sequences in common between paralogous *Hox* clusters. ParaCNS1, paraCNS2, and paraCNS3 are all located immediately (approximately 2 kb/approximately 4 kb) downstream of *EVX1/EVX2* (Fig. 1, red asterisks) and are within the previously described putative regulatory element R1 (Goodman et al. 2002). ParaCNS4 and paraCNS5 are located within 1 kb of each other in both the context of 5′ *HOXD13* and 5′ *HOXA13* but are individually situated over 200 kb away from *HOXD13* and over 350 kb away from *HOXA13* (Fig. 1, blue asterisks).

ParaCNS4 and paraCNS5 are core identities of 131 nt/68% and 105 nt/60%, respectively, yet when these core sequences were compared with homologous sequences of other species, they fell within larger highly conserved sequences. 5′ of *HOXA13*, paraCNS4 and paraCNS5 fall within the fourth intron of *HIBADH* embedded within a CNS of approximately 2.25 kb: A13CNS (Fig. 2). These two paraCNSs also fall within an approximately 3-kb CNS 5′ of *HOXD13*: D13CNS. Notably, they are otherwise unique in the mouse and human genomes. Interestingly, D13CNS is a candidate conserved sequence within the previously identified critical region for the "distal limb enhancer" (Spitz et al. 2003). The coincidence of the paralogous conservation with the distal limb enhancer critical region makes its paralogous counterpart, A13CNS, an appealing enhancer candidate for global regulation of *HOXA13* and 5′ genes.

A13CNS function was directly tested in embryonic development in transgenic mice. Mouse A13CNS (mmA13CNS) was placed in *cis* with a minimal *Hsp68-lacZ* construct and used as a transgene for E12.5 founder analysis of the embryos. *LacZ* staining was observed in a variety of developing structures (Fig. 3). Staining was seen in the superior neural layer and wall, but not the overlying epithelium, of the midbrain (Fig. 3D). In addition, *lacZ* staining was observed in the presumptive olfactory tissue (Fig. 3E), hindbrain, pons, medulla, and corpus plexus (Fig. 3F). Staining was observed in the distal mantle region of the lumbosacral spinal cord, midway

Table 1. Characterization of *HoxA/D* paraCNSs

Conserved Element	Distance from *HOXD13*	Distance from *HOXA13*	Length of Homology	% Identity
paraCNS1	16,688 bp	48,674 bp	209 bp	64
paraCNS2	17,026 bp	49,509 bp	47 bp	87
paraCNS3	19,042 bp	51,049 bp	205 bp	64
paraCNS4	238,369 bp	351,994 bp	131 bp	68
paraCNS5	239,106 bp	353,107 bp	105 bp	66

HsA13CNS

Ptr_A (Chimp) — 100%
— 50%
Cfa_A (Dog)
Mmu_A (Mouse)
Rno_A (Rat)
Gga_A (Chick)
Fru_A (Fugu)
Hsa_D (Human)
Mmu_D (Mouse)

* *
P4   P5   1kb

```
Hsa_A  AAGTAATCAACAAATTTAAGGGTGTTAATTAACTTTACACTTTA
Ptr_A  ...........................................
Cfa_A  ...........................................
Mmu_A  ...........................................
Rno_A  ...........................................
Gga_A  ...........................................
Fru_A  .................A.G.G.......................
Hsa_D  .G.C...T.....C.C.T...........T...C.T...G..
Mmu_D  .G.C...T.....C.C.T...........T...C.T...G..

Hsa_A  TCATGTTCAATTAG-TTGTCCTTTCTAGAAAAGATTGCCGTGCA
Ptr_A  ..............-............................
Cfa_A  ..............-............................
Mmu_A  ..............-............................
Rno_A  ..............-............................
Gga_A  ...................................C.......
Fru_A  ..............-GCA...C.C.------------.A...-
Hsa_D  ..............C..AGT..C.GC.AC....GCAA.GA..A.
Mmu_D  ..............C..AGT..C.GC.AC....GC.A.GA..A.

Hsa_A  CAATCAGTTTA--GAGCGTGCCATGTTTGATTTGTAAGTGATC
Ptr_A  ...........--..............................
Cfa_A  ...........--..............................
Mmu_A  ...........--........................C.....
Rno_A  ...........--..............................
Gga_A  .....G.....--........................C.....
Fru_A  .TGCTG.C...AG..............................
Hsa_D  ....TGC...GATAG....A.AG.AC..C.......GT.A...
Mmu_D  ....TGC...GATAG....A.AG.AC..C.......GT.A...
```
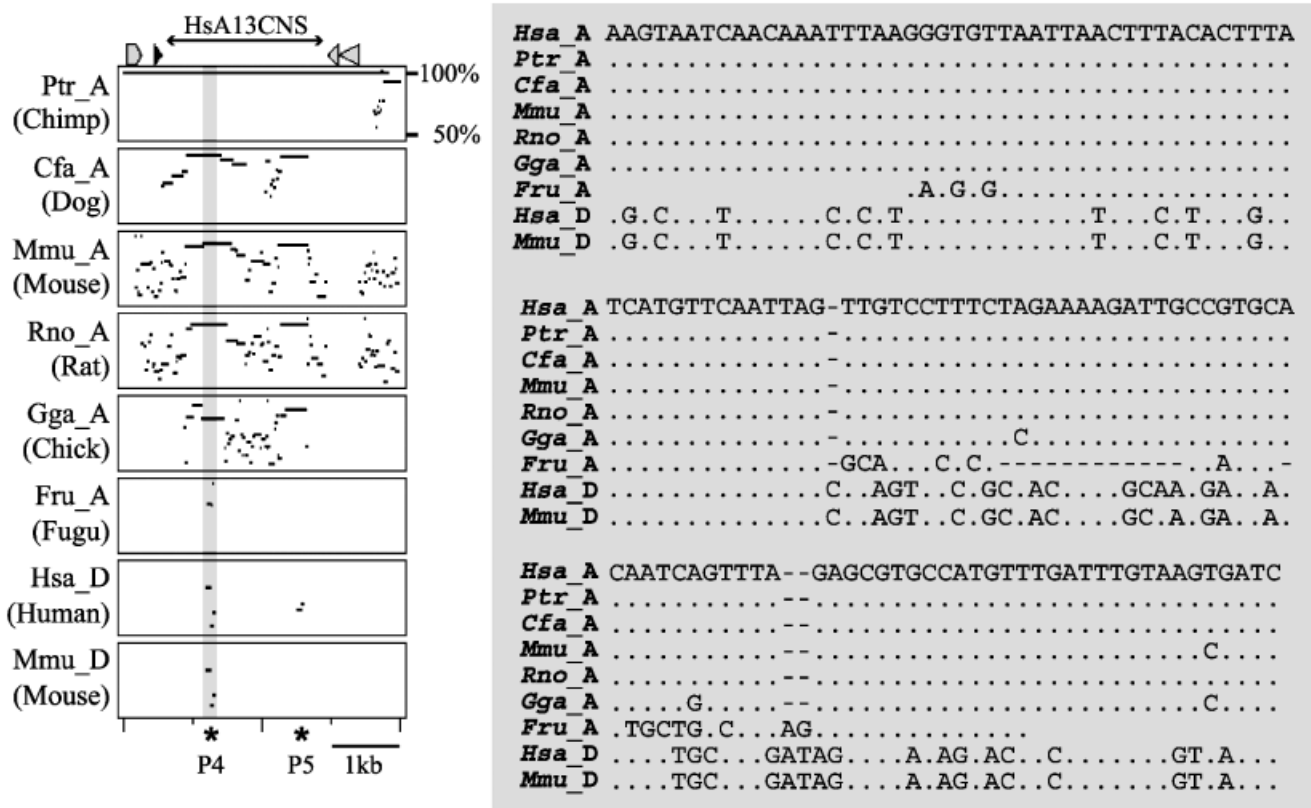
Fig. 2. Evolutionary conservation of HsA13CNS. Multipipmaker was used to align human A13CNS with six orthologous regions (Hsa, *Homo sapiens*; Ptr, *Pan troglodytes*; Cfa, *Canis familiaris*; Mmu, *Mus musculus*; Rno, *Rattus norvegicus*; Gga, *Gallus gallus*; Fru, *Takifugu rubripes*) and the human and mouse paralogous regions 5′ of *HOXD13*. Length of sequence identities versus the human reference sequence are depicted as horizontal lines, and the percent identity from 50 to 100 is measured on the vertical axis. Asterisks represent location of paraCNS4 and paraCNS5. Sequence alignment on the right within vertical gray box is paraCNS4 depicted at the nucleotide level, where (·) represents identity.

between the ventral and dorsal gray horns of the spinal cord (Fig. 3D), and weak staining was visualized in a portion of the mesenchyme of the genital tubercle and the anterior wall of the urogenital sinus (Fig. 3G). Additionally, variable weak staining was observed in the interdigital mesenchyme and proximal limb regions in 5 of 12 expressing founder limbs (Fig. 4, A and B).

Two permanent lines of mice were established for the mmA13CNS-*lacZ* transgene. Analysis of transgenic embryos from one line at E11.5–E13.5 yielded strong central nervous system staining identical to that of the founders. The neural tube, midbrain, and hindbrain lacZ expression seen in the mmA13CNS mouse is highly similar to the neural expression patterns of *Evx1* (neural tube and hindbrain) and *Evx2* (neural tube, hindbrain, and midbrain) (Dollé et al. 1994) (Fig. 5). Limb staining was observed in proximal limb elements and the presumptive digits at E12.5 and E14.5 (Fig. 4). The lacZ staining in the proximal limb elements at these times (Fig. 4, C–E) may represent *HoxA* expression in the developing limb musculature (Yamamoto et al. 1998). However, further

analysis involving myoblast markers must be used to definitively characterize this lacZ expression pattern, but we hypothesize that a long-range enhancer of posterior *HoxA* genes in muscle development may lie within mmA13CNS. Profound staining was seen at E13.5 and E14.5 in the olfactory tissue or developing vomeronasal organ (data not shown). Genital mesenchyme lacZ expression that was seen in three of three expressing founders was not seen in any offspring of this permanent line, suggesting a potential negative effect of the transgene insertion site.

In the second permanent line, identical central nervous system expression was observed. Staining in offspring of this line at E11.5 revealed limb and genital bud lacZ staining; however, as with the other line, the precise *Hoxa13* expression pattern was not observed. Thus, although timing and general tissue expression may be partly recapitulated with this construct, precise domain expression was not observed. The wide range of reproducible lacZ expression domains confirms the functional capabilities of this highly conserved CNS. The observation of broader reproducible expression in tissues not
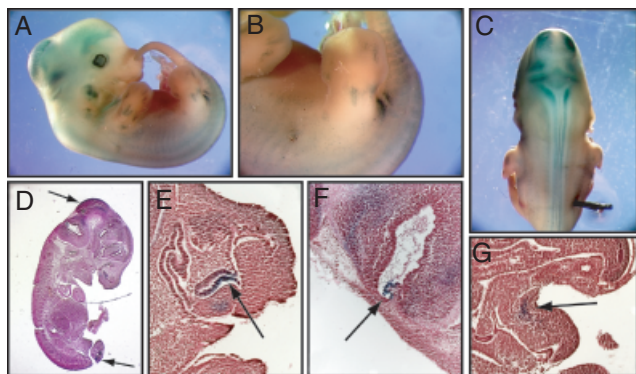
Fig. 3. LacZ expression of mmA13CNS founders at E12.5. Representative *lacZ* staining in one of three E12.5 transgenic founders. (A) Wide range of lacZ-positive structures. (B) Whole-mount expression in genital bud in addition to weak limb staining. (C) Central nervous system staining. (D–G) Ten-micrometer paraffin sections of the founder in A–C, with standard hematoxylin and eosin staining. (D) LacZ expression specific to the periphery of the midbrain (top arrow) and caudal neural tube (bottom arrow). (E) Staining specific to the olfactory region. (F) The expression found in the corpus plexus (arrow) and within the hindbrain. (G) Staining in the genital bud mesenchyme.

known to express any of the upstream neighboring genes suggests mmA13CNS may harbor enhancer elements, perhaps for distant genes, capable of directing expression to those tissues. Alternatively, essential negative control elements may be missing. Finally, this 2.25-kb element is not sufficient for full recapitulation of limb bud and genital bud expression for this domain of genes.
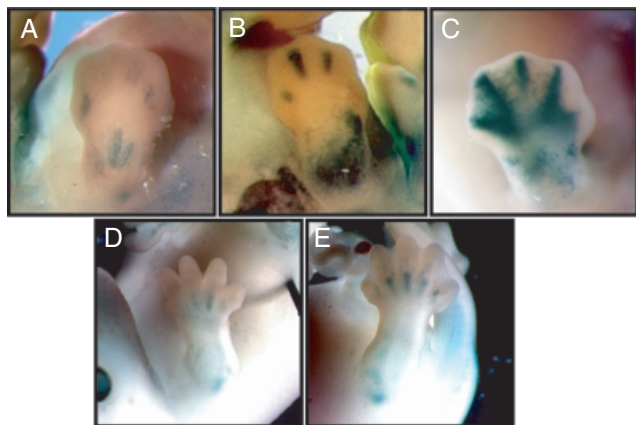

Fig. 4. LacZ expression in mmA13CNS transgenic limbs. LacZ-stained E12.5 (A–C) and E14.5 (D and E) embryos. (A and B) mmA13CNS transgenic founders. Staining is shown in the inter-digital mesenchyme and in proximal limb elements. (C–E) Obtained from the mmA13CNS permanent line. (C) Digital lacZ staining and a diffuse staining in the autopod and the proximal limb. Representative forelimb (D) and hindlimb (E), with distinct digital and proximal limb staining.
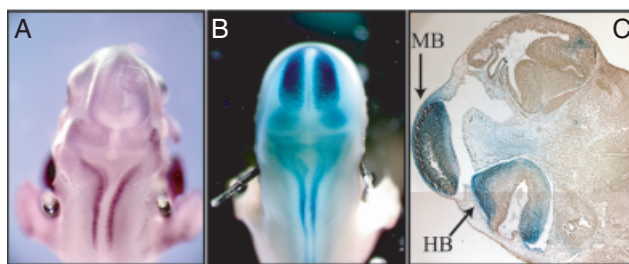

Fig. 5. *Evx1* versus mmA13CNS expression. (A) *Evx1* RNA neural tube staining at E12.5. (B) Comparable neural tube staining (lacZ) in the mmA13CNS transgenic E12.5 embryo. (C) Six-micrometer paraffin section of the lacZ-stained mmA13CNS transgenic embryo in B with eosin counterstain. The hindbrain (HB) is a known site of expression for both *Evx1* and *Evx2*,

DISCUSSION

A global regulatory mechanism affects not only *Hoxd* but also *Hoxa*

Previous reports showed evidence for multiple long-range global regulatory elements 5′ of the *HoxD* cluster (Spitz et al. 2001). It has been proposed that distinct long-range enhancer elements are capable of driving neural, distal limb, and genital bud specific expression of genes within a fixed genomic region surrounding *Hoxd13* from E10.5 through E13.5. The distal limb enhancer is the most finely mapped of these elements, being localized to a 54-kb region approximately 200 kb 5′ of mouse *Hoxd13*. Spitz et al. (2003) showed that this region is capable of driving distal limb expression similar to that of the *AbdB*-like *HoxD* genes in two non-*Hox* genes: *Evx2* and the newly identified *Lnp* gene. A 5′ boundary to the influence of the distal limb enhancer cannot be determined because of the dearth of genes in the region; however, this study revealed that a long-range enhancer potentially activates all genes within a finite genomic domain rather than selected promoters.

Characterizing the *Hoxa13* gene neighborhood for expression during E10.5 through E13.5 revealed four consecutively arranged genes in addition to *HOXA13* whose expression was temporally and spatially similar, suggesting the influence of a regulatory mechanism that promotes distal limb and genital bud expression in disparate genes within a general domain. The genes within this region have an expression pattern highly similar to those under the control of the distal limb enhancer 5′ of the *HoxD* cluster. In addition, the fact that mmA13CNS can drive an *Evx1*-like neural expression pattern and is paralogous to mmD13CNS which is 5′ of *Evx2* shows that the element, and probably the regulatory effect on *Evx* expression in the central nervous system, existed before the divergence of the *Hox* clusters. Altogether, it follows that the *HoxA* and *HoxD* regulatory mechanisms have been conserved since

duplication rather than arising independently by convergent evolution.

The expression profiles of limb/genital bud expression of *Hoxa13*, *Evx1*, *Hibadh*, *Tax1bp1*, *Jaz1*, and *Crebpa* at E11.5 define a finite regulatory domain for the distal limb enhancer (Fig. 1). Based on the presence or absence of distal limb expression at E11.5, the domain in human extends across 1,227,381 bp from, but not including, *HOXA11* to the 5′ end of *CREB5*. This presumes that the ubiquitous expression, including distal limb and genital bud, of *Jaz1* at this stage obscures these discrete expression domains. Based on the distal limb expression of these genes seen at E12.5 and later, it might be argued that the domain of regulation may include the promoter of *CREB5*. If so, the domain may be larger and lack of distal limb expression of *Crebpa* at E11.5 may be related to the relatively longer distance of the promoter to the regulatory element. This hypothesis might also explain the lack of tight distal autopod restriction seen with *Tax1bp1* and *Jaz1* at E11.5.

Is there a role for the genes upstream of *Hoxa13* to be expressed in the developing limbs and genital bud? The regulatory domains that exist upstream of the *HoxA* and *HoxD* clusters provide an intriguing problem of genome evolution in relation to long-range enhancers. Distinctly different genes occupy the regions adjacent to mmA13CNS and mmD13CNS; therefore, substantial gene insertions and/or deletions occurred near these extant conserved enhancer elements. In the context of insertions, such elements could have conferred their regulatory capabilities and restraints to the newly introduced genes. It follows that some genes expressed in new embryonic domains could be deleterious and might not be maintained. Alternatively, new gene addition(s), and expression(s), could have no effect or could lead to a selective advantage. As a result, genes having no prior role in limb or genital development could influence patterning merely by being transposed into a regulatory domain. At this time it is not known how the current genomic structure evolved or whether the genes upstream of *Hoxa13* and *Hoxd13* have a role in limb or genital development.

Paralogous cluster noncoding sequence conservation

It has been established that exonic/coding nucleotide sequences tend to be conserved between evolutionarily distant species, such as human and mouse, based on the functional constraints on protein coding or mRNA splicing domains. It is becoming increasingly evident that certain noncoding sequences are also similarly highly conserved. These CNSs have been associated with functional regulatory elements (Loots et al. 2000), and CNSs may act over large (>1 Mb) genomic distances (Lettice et al. 2003). Analysis of large genomic regions for potential long-range regulatory elements by com-

parative sequence analysis between human and mouse homologous sequences results in a large number of CNSs. For example, 99 non-exonic CNSs greater than 100 nucleotides in length and an identity greater than 80% were found in a comparison of over 900 kb 5′ of *HOXD13* versus the homologous mouse region. Similarly, comparing over 1 Mb of human genomic sequence 5′ of *HOXA13* with the homologous mouse sequence yielded 138 CNSs longer than 100 nucleotides with greater than 80% identity. This is not surprising given the relatively close evolutionary relationship between mouse and human. Yet, the prospect of systematically analyzing these conserved sequences for their developmental functional capacity in transgenic mice is daunting. Potential improvements in CNS selection might be made by inclusion of more distant species in the sequence analysis (Margulies et al. 2003) or by paralogous comparisons.

If the rudiments of a conserved regulatory network existed before Hox cluster divergence and remained operative, then paralogous sequence comparison should allow identification of ancient regulatory elements in common with both the *HoxA* and *HoxD* clusters that are essential for this mechanism. The fact that the 2.25-kb mmA13CNS is so highly conserved among osteichthyans is highly suggestive of a critically important function. Because this conservation can be extended to the paralogous mmD13CNS, which is already known to be within the critical region for the *HoxD* distal limb enhancer, the case for mmA13CNS as a regulatory element(s) becomes stronger. Yet, the fact that this conservation is observed as far back as fugu within paraCNS4 (Fig. 2) seems counterintuitive because of the insinuated function of the distal limb enhancer in digit formation. Evidence shows that posterior *HoxA* orthologs in zebrafish also exhibit distal fin expression (Géraudie and Birraux 2003). Thus, the distal limb enhancer may not function per se in digit formation but merely may promote general domain and time of gene expression in appendage development, whether it is distal limb or distal fin. Alternatively, this regulatory element may have existed first to promote genital development, with involvement in limb development evolving secondarily.

A 54-kb critical region 5′ of *HOXD13* is capable of directing distal limb mesenchymal expression of *Hoxd13*, and mmD13CNS is located entirely within this critical region. Its paralog, mmA13CNS, by itself as a 2.25-kb segment, does not drive reporter expression in the entire distal limb domain nor in the genital bud as expected. Yet, this sequence is sufficient for *Evx*-like reporter expression in the central nervous system. LacZ recombinant BAC constructs spanning mmA13CNS in its endogenous genomic context may help to elucidate whether additional sequences will allow for faithful distal limb and genital bud expression. If successful, an mmA13CNS-*lacZ*-deleted BAC transgene would directly test its necessity in limb and genital bud development.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215: 403–410.

Bastian, H., and Gruss, P. 1990. A murine *even-skipped* homologue, *Evx 1*, is expressed during embryogenesis and neurogenesis in a biphasic manner. *EMBO J.* 9: 1839–1852.

Bober, E., Franz, T., Hans-Henning, A., Gruss, P., and Tremblay, P. 1994. *Pax-3* is required for the development of limb muscles: a possible role for the migration of dermomyotomal muscle progenitor cells. *Development* 120: 603–612.

De Valck, D., et al. 1999. The zinc finger protein A20 interacts with a novel anti-apoptotic protein which is cleaved by specific caspases. *Oncogene* 18: 4182–4190.

DiLeone, R. J., Russell, L. B., and Kingsley, D. M. 1998. An extensive 3′ regulatory region controls expression of Bmp5 in specific anatomical structures of the mouse embryo. *Genetics* 148: 401–408.

Dollé, P., Fraulob, V., and Duboule, D. 1994. Developmental expression of the mouse *Evx-2* gene: relationship with the evolution of the HOM/Hox complex. *Development* (suppl.): 143–153.

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

Ferrier, D. E. K., Minguillón, C., Holland, P. W. H., and Garcia-Fernàndez, J. 2000. The amphioxus Hox cluster: deuterostome posterior flexibility and *Hox14*. *Evol. Dev.* 2: 284–293.

Fromental-Rain, C., Warot, X., Messadecq, N., Lemeur, M., Dolle, P., and Chambon, P. 1996. *Hoxa-13* and *Hoxd-13* play a crucial role in the patterning of the limb autopod. *Development* 122: 2997–3011.

Géraudie, J., and Birraux, V. B. 2003. Posterior *hoxa* genes expression during zebrafish bony fin ray development and regeneration suggests their involvement in scleroblast differentiation. *Dev. Genes Evol.* 213: 182–186.

Goodman, F. R., et al. 2000. Novel HOXA13 mutations and the phenotypic spectrum of hand-foot-genital syndrome. *Am. J. Hum. Genet.* 67: 197–202.

Goodman, F. R., Majewski, F., Collins, A. L., and Scrambler, P. J. 2002. A 117-kb microdeletion removing *HOXD9–HOXD13* and *EVX2* causes synpolydactyly. *Am. J. Hum. Genet.* 70: 547–555.

Hogan, B., Beddington, R., Costantini, F., and Lacy, E. 1994. *Manipulating the Mouse Embryo—A Laboratory Manual.* 2nd Ed. Cold Spring Harbor Laboratory Press, Plainview, NY.

Kondo, T., Zakany, J., Innis, J. W., and Duboule, D. 1997. Of fingers, toes, and penises. *Nature* 390: 29.

Koontz, J. I., et al. 2001. Frequent fusion of the JAZF1 and JJAZ1 genes in endometrial stromal tumors. *Proc. Natl. Acad. Sci. USA* 98: 6348–6353.

Kothary, R., Clapoff, S., Darling, S., Perry, M. D., Moran, L. A., and Rossant, J. 1989. Inducible expression of an *hsp68-lacZ* hybrid gene in transgenic mice. *Development* 105: 707–714.

Lettice, L. A., et al. 2003. A long-range *Shh* enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum. Mol. Genet.* 12: 1725–1735.

Loots, G. G., et al. 2000. Regulator of interleukins 4, 13, and 5 by cross-species sequence comparisons. *Science* 288: 136–140.

Margulies, E. H., Blanchette, M., Haussler, D., and Green, E. D.; NISC Comparative Sequencing Program. 2003. Identification and characterization of multi-species conserved sequences. *Genome Res.* 13: 2507–2518.

McGinnis, W., and Krumlauf, R. 1992. Homeobox genes and axial patterning. *Cell* 68: 283–302.

Mortlock, D. P., and Innis, J. W. 1997. Mutation of HOXA13 in hand-foot-genital syndrome. *Nat. Genet.* 15: 179–180.

Mortlock, D. P., Post, L. C., and Innis, J. W. 1996. The molecular basis of *Hypodactyly* (*Hd*): a deletion in *Hoxa13* leads to arrest of digital arch formation. *Nat. Genet.* 13: 284–289.

Post, L. C., and Innis, J. W. 1999. Altered Hox expression and increased cell death distinguish Hypodactyly from *Hoxa13* null mice. *Int. J. Dev. Biol.* 43: 287–294.

Rougraff, P. M., Paxton, R., Kuntz, M. J., Crabb, D. W., and Harris, R. A. 1988. Purification and characterization of 3-hydroxyisobutyrate dehydrogenase from rabbit liver. *J. Biol. Chem.* 263: 327–331.

Ruddle, F. H., Bartels, J. L., Bentley, K. L., Kappen, C., Murtha, M. T., and Pendleton, J. W. 1994. Evolution of *Hox* genes. *Annu. Rev. Genet.* 28: 423–442.

Schwartz, S., et al. 2000. Pipmaker—a web server for aligning two genomic DNA sequences. *Genome Res.* 10: 577–586.

Spitz, F., Gonzalez, F., and Duboule, D. 2003. A global control region defines a chromosomal regulatory landscape containing the *HoxD* cluster. *Cell* 113: 405–417.

Spitz, F., Gonzalez, F., Peichel, C., Vogt, T. F., Duboule, D., and Zákány, J. 2001. Large scale transgenic and cluster deletion analysis of the *HoxD* complex separate an ancestral regulatory module from evolutionary innovations. *Genes Dev.* 15: 2209–2214.

Stadler, H. S., Higgins, K. M., and Capecchi, M. R. 2001. Loss of *Eph-receptor* expression correlates with loss of cell adhesion and chondrogenic capacity in *Hoxa13* mutant limbs. *Development* 128: 4177–4188.

Veraksa, A., Del Campo, M., and McGinnis, W. 2000. Developmental patterning genes and their conserved functions: from model organisms to humans. *Mol. Genet. Metab.* 69: 85–100.

Warot, X., Fromental-Rain, C., Fraulob, V., Chambon, P., and Dolle, P. 1997. Gene dosage-dependent effects of the *Hoxa-13* and *Hoxd-13* mutations on morphogenesis of the terminal parts of the digestive and urogenital tracts. *Development* 124: 4781–4791.

Yamamoto, M., et al. 1998. Coordinated expression of *Hoxa-11* and *Hoxa-13* during limb muscle patterning. *Development* 125: 1325–1335.

Zakany, J., Fromental-Ramain, C., Warot, X., and Duboule, D. 1997. Regulation of number and size of digits by posterior Hox genes: a dose-dependent mechanism with potential evolutionary implications. *Proc. Natl. Acad. Sci. USA* 94: 13695–13700.