Mar. Biotechnol. 2, 231–240, 2000
DOI: 10.1007/s101269900027

MARINE BIOTECHNOLOGY

Identification of *Hox* Gene Sequences in the Sea Cucumber *Holothuria glaberrima* Selenka (Holothuroidea: Echinodermata)

Ana T. Méndez, José L. Roig-López, Pedro Santiago, Carlos Santiago, and José E. García-Arrarás*

Department of Biology, University of Puerto Rico, Rio Piedras, P.R. 00931

Abstract: The Echinodermata is a unique animal group forming an early branch in the deuterostomes phylogenetic tree. In echinoids and asteroids a single *Hox* cluster with nine cognates of the vertebrate *Hox* paralogous groups has been reported, but no data are available from other echinoderm classes. We report here nine *Hox*-type sequences from the sea cucumber *Holothuria glaberrima,* a member of the class Holothuroidea. Partial homeodomain sequences were amplified by polymerase chain reaction from genomic DNA and from a regenerating gastrointestinal tract complementary DNA library. Sequence analyses suggest that the holothuroid cluster has at least three genes of the anterior, one of the medial, and five of the posterior groups. This is the first evidence of five posterior sequences in echinoderms.

Key words: Deuterostomes, echinoderms, holothuroids, paralogous group, partial homeodomain, orthologues

INTRODUCTION

Homeobox-containing genes are a family of regulatory genes involved in pattern formation and cell fate determination (reviewed in Akam, 1989; Gehring, 1992). They have a highly conserved 180-bp sequence, the homeobox, encoding a 60 amino acid homeodomain with a DNA-binding helix-turn-helix motif. Homeodomain proteins recognize and specifically bind to DNA responsive elements to control the transcription of genes mediating the cellular processes dictating morphogenesis (as reviewed by Carroll, 1995).

In *Drosophila,* there are eight *Hox*-type homeobox genes within the homeotic complex called HOM-C. These genes share a high degree of sequence similarity with the *Antennapedia* (*Antp*) gene homeodomain, and in embryogenesis their expression domains along the anterior-to-posterior axis follow the same order as the arrangement of genes along the chromosome (reviewed in McGinnis and Krumlauf, 1992; Carroll, 1995; Holland and García-Fernàndez, 1996). *Hox*-type genes in other metazoans share a highly conserved homeodomain sequence, cluster within chromosomes, and express in discrete regions during development. It has been proposed that with the exception of sponges and cnidarians, most invertebrate metazoans have a single *Hox* complex while vertebrates have four (Balavoine, 1998; Holland, 1998; Martinez et al., 1998). Exceptions to this rule are the horseshoe crab *Limulus polyphemus,* in which four *Hox* clusters might be found (Cartwright et al., 1993), and teleost, in which up to seven *Hox* clusters have been documented (Amores et al., 1998). The

Received July 19, 1999; accepted October 18, 1999.

*Corresponding author: University of Puerto Rico, Río Piedras, Box 23360, San Juan, PR 00931-3360; telephone 787-764-0000, ext 2596; fax 787-764-3875; e-mail jegarcia@upracd.upr.clu.edu

Hox complexes in vertebrates are found in different chromosomes and contain 9 to 11 genes that can be aligned into 13 sets of paralogous groups (Scott, 1993; Holland and García-Fernàndez, 1996). Not all genes are found in each cluster, and only eight of them are orthologous (cognate) to the ones found in *Drosophila,* with paralogous groups 9–13 thought to be duplicate copies of *Drosophila's Abdominal-B* gene (Carroll, 1995). The expression pattern of *Hox* genes during development and homeodomain sequence comparisons indicate a relatedness between paralogous groups 1–3, 4–8, and 9–13, which have been termed as the anterior, medial, and posterior groups, respectively (Ruddle et al., 1994a).

In echinoderms, most of the work has been carried out in sea urchins (Echinoidea), in which 10 *Hox* genes have been reported, and proposed to be part of a single echinoid *Hox* cluster (Popodi et al., 1996; Martinez et al., 1999). Sea urchin *Hox*-type genes were named by the species initials followed by "Hbox" with a number that indicates the order in which they were reported rather than the cognate group homology (Scott, 1993; Ruddle et al., 1994b; Popodi et al., 1996; Morris et al., 1997). Analysis of sea urchin homeodomain sequences served to assign them within the anterior, medial and posterior paralogy groups (Ruddle et al., 1994a; Popodi et al., 1996; Morris et al., 1997). The posterior group sequences are *Tripneustes gratilla* Hbox4 (Dolecki et al., 1988), *Strongylocentrotus purpuratus* Hbox7 (Zhao, 1992; Dobias et al., 1996), and, *Lytechinus variegatus* and *S. purpuratus* Hbox10 (Ruddle et al., 1994a). The medial group sequences are *S. purpuratus* and *L. variegatus* Hbox9 (Ruddle et al., 1994a) and *T. gratilla* Hbox3, Hbox6, and Hbox1 (Dolecki et al., 1986, 1988; Wang et al., 1990). Until recently, the only anterior sequence that had been reported in sea urchins was *Paracentrotus lividus* Hbox11 (Di Bernardo et al., 1994). However, recent work by Martinez and colleagues (1999) shows the presence of paralog genes to *Hox1* and *Hox2* groups from *S. purpuratus.* They have also shown that the echinoid *Hox* genes are organized in a single cluster and that this cluster is essentially similar to that in other chordates.

Furthermore, seven *Hox*-type sequences were isolated from the sea star (Asteroidea) *Asterina minor* (Mito and Endo, 1997). These sequences were named by the species initials followed by a number assigned in relation to their possible anterior, medial, and posterior paralogy grouping. The *A. minor* sequences are AM-1, AM-3, AM-4, AM-5, AM-6, AM-7, and AM-9. Partial homeodomain sequence comparison suggests that AM-1 and AM-3 are derived from the anterior cognate groups, AM-4 to AM-7 from the medial cognate groups, and AM-9 from the posterior cognate groups.

Thus, the data available to date from the sea urchins and sea stars define a *Hox* cluster with at least 10 genes within the phylum Echinodermata. However, the possibility of other cognates within the described cluster has not been ruled out (Popodi et al., 1996). Furthermore, there is a lack of data on *Hox*-type genes found in other classes of echinoderms—namely, the class Ophiuroidea (brittle stars), class Crinoidea (sea lilies and feather stars), and class Holothuroidea (sea cucumbers). The latter is particularly interesting because of its derived secondary bilateral symmetry within the otherwise radial phylum.

Therefore, we have employed the polymerase chain reaction (PCR) with degenerate primers (Pendleton et al., 1993; Mito and Endo, 1997) to isolate *Hox*-type sequences from the sea cucumber *Holothuria glaberrima* Selenka. The experimental strategy has been used successfully by several investigators (Murtha et al., 1991; Pendleton et al., 1993; Ruddle et al., 1994a; Degnan et al., 1995; Tarabykin et al., 1995; Washabaugh et al., 1995; Bayascas et al., 1997; Finnerty and Martindale, 1997; Mito and Endo, 1997; Morris et al., 1997). Nine sea cucumber *Hox* sequences were isolated from genomic DNA and a cDNA library of regenerating gastrointestinal tract: HgHbox1, HgHbox2, HgHbox3, HgHbox5, HgHbox9, HgHbox10, HgHbox11, HgHbox12, and HgHbox13. Provisional orthologue assignments were deduced by phylogenetic analysis and partial homeodomain similarity comparisons between taxa (Ruddle et al., 1994a; Popodi et al., 1996; Mito and Endo, 1997; Morris et al., 1997).

Materials and Methods

Genomic DNA Isolation

Genomic DNA (gDNA) was prepared from the gonads of the sea cucumber *Holothuria glaberrima.* Animals (10–15 cm) were collected from the rocky shores of the north side of Puerto Rico and kept in seawater aquaria (20°C to 24°C) until dissections were performed. Sea cucumbers were anesthetized by immersion in ice cold water for 1 hour. Gonads were excised through a longitudinal slit of the body wall, placed immediately in liquid nitrogen, and subsequently

ground to a powder. Up to 100 mg of gonad powder was diluted in 3 ml of 50 mM Tris-HCl (pH 9.0), 100 mM EDTA (pH 8.0), 200 mM NaCl, 0.4% sodium dodecyl sulfate (SDS), and 200 mg/ml proteinase K, and incubated while being gently shaken at 55°C for 3 hours (Sambrook et al., 1989). The solution was extracted several times with phenol and once with chloroform/isoamyl alcohol (24:1, vol/vol). The DNA was purified by precipitation with 0.1 volume of 3.0 M sodium acetate (pH 5.0) and 2 volumes of 95% ethanol, centrifuged, and resuspended in water.

RNA Isolation and cDNA Library Preparation

Regenerating gastrointestinal tracts (5 to 7 days after evisceration or regeneration stage II; García-Arrarás et al., 1998) from *H. glaberrima* sea cucumbers were used for total RNA isolation. The single-step method was followed using the acid guanidinium thiocyanate phenol chloroform extraction procedure (Chomczynski and Sacchi, 1987). Briefly, 0.5 g of tissue was homogenized in 5 ml of a solution consisting of 4 M guanidinium thiocyanate, 25 mM sodium citrate (pH 7.0), 0.5% sarcosyl, 100 mM 2-mercaptoethanol, 3 M sodium acetate (pH 4.0), and 5 ml of water-saturated phenol. After homogenization, 1 ml of chloroform was added to the homogenate, followed by vigorous vortex mixing. The homogenate was centrifuged and the aqueous phase was precipitated overnight with an equal volume of isopropanol. Total RNA was resuspended in ultrapure RNAse-free water. Poly(A)$^+$ RNA was purified by oligo(dT)-cellulose column chromatography (Aviv and Leder, 1972) according to standard methods (Sambrock et al., 1989). One milligram of total RNA was loaded into an oligo(dT)-cellulose packed column. The column was first washed with binding buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA, and 0.5 M NaCl), and then washed with low-salt buffer (10 mM Tris-Cl, pH 7.5, 1 mM EDTA, and 0.1 M NaCl). The poly(A)$^+$ RNA was eluted from the column with a solution consisting of 10 mM Tris-Cl (pH 7.6), 1 mM EDTA. Further enrichment of mRNA was obtained by passing the eluate a second time through the column. The eluted poly(A)$^+$ RNA was precipitated with 2 volumes of 95% ethanol and 0.1 M NaCl, centrifuged, and resuspended in ultrapure RNAse-free water.

The cDNA library was constructed using the Uni-Zap XR vector and the Zap-cDNA synthesis kit following the manufacturer's instructions (Stratagene). The titer of the unamplified cDNA library was 4.2×10^5 pfu/ml and 2.15×10^9 pfu/ml after amplification. The percentage of recombinants was 89.5% and the range of insert size was 0.5 to 3.0 kb (based on the analysis of 38 inserts).

PCR Amplifications

PCR amplifications of gDNA and the cDNA library from regenerating gastrointestinal tracts were done with the degenerate primer sets *Hox* E and *Hox* F (Pendleton et al., 1993), and 27-mer and 40-mer (Mito and Endo, 1997), respectively. The primer sequences were: *Hox* E, 5′-AAAGGATCCTGCAGARYTIGARAARGARTT-3′; *Hox* F, 5′-ACAAGCTTGAATTCATICKICKRTTYTGRAACCA-3′; 27-mer, 5′-AAAAGGATCCGARCTNGARAARGARTT-3′; and 40-mer, 5′-AAAACTGCAGYTTCATNCGNCGRTTYTGRAACCADATYTT-3′, where R = A or G, Y = C or T, K = G or T, I = inosine, D = A, G, or T, and N = A, C, G, or T. These primers match highly conserved regions of the *Hox*-type gene homeobox. *Hox* E and *Hox* F generate a fragment of 82 bp (27 amino acids) in length corresponding to amino acid positions 21 to 47 of the homeobox. The 27-mer and 40-mer primers generate a fragment of 76 bp (25 amino acids) in length spanning amino acid positions 21 to 45 of the homeobox. Each 100-μl PCR reaction included 1.0 μg of gDNA or 10 μl of the amplified cDNA library as template, 2.5 U of *Taq* polymerase (Promega) in the supplier's buffer, 200 μM of each dNTP, 1.5 mM $MgCl_2$, and 50 pmol of each primer. For the library PCR the 10 μl of template cDNA was first mixed with ultrapure water and warmed for 5 minutes at 70°C. The PCR reaction mixture (described above) was then added. The reactions were performed in a Robocycler 40 temperature thermocycler (Stratagene) with a hot start: 95°C for 5 minutes followed by addition of 2.5 U of *Taq* DNA polymerase, and then 2 cycles of 97°C for 1 minute, 40°C for 1 minute, 72°C for 1 minute; followed by 38 cycles of 94°C for 50 seconds, 40°C for 1 minute, 72°C for 1 minute; incubation at 72°C for 5 minutes (as recommended by Mito and Endo, 1997).

In addition, PCR reactions with the same primers were performed to verify the cloned fragment size within the recombinant colony (colony PCR). In the colony PCR reactions, an inoculate of the recombinant positive colony (38 colonies from genomic DNA and 72 colonies from the cDNA library) was diluted in 50 μl of cell lysis buffer (1% Triton X-100, 2 mM EDTA, and 20 mM Tris, pH 8), and warmed for 5 minutes at 95°C. Ten microliters of this dilution was used as template under the same PCR reaction

mixture conditions mentioned above. The PCR program was one cycle of 94°C for 1 minute, 28 cycles at 94°C for 1 minute, 55°C for 2 minutes, and 72°C for 3 minutes, with a final incubation at 72°C for 5 minutes.

Cloning and Sequencing

PCR products were analyzed by agarose gel electrophoresis and cloned into a pCR 2.1 plasmid using the Original TA-Cloning kit (Invitrogen) following the manufacturer's instructions. This was done by taking 1 μl of the PCR reaction product and ligating it into the TA cloning vector. Recombinant plasmids were then transformed into *Escherichia coli* Top10F′ competent cells (Invitrogen). Recombinant clones were selected by blue/white screening and by colony PCR reactions (see "PCR Amplifications"). Clones with inserts of 130–150 bp were grown for sequence analysis. Purified plasmid DNA from recombinant clones was sequenced in forward and reverse by the Sanger dideoxy method (Sanger et al., 1977) using the Cy5 AutoRead Sequencing kit and the ALFexpress DNA Sequencer (Pharmacia Biotech).

Sequence Comparisons

The informative 82-bp and 76-bp sequences from *H. glaberrima* were translated (ORF Finder) and compared against the databases accessible through GenBank. Nucleotide and amino acid alignment studies were performed using the BCR Search Launcher followed by multiple sequence alignments with the CLUSTALW 1.7 Multiple Sequence Alignments database (Human Genome Center, Baylor College of Medicine, Houston, Tex.). The *H. glaberrima Hox*-like sequences were assigned to cognate groups on the basis of amino acid similarity to echinoderm sequences (Popodi et al., 1996; Mito and Endo, 1997; Martinez et al., 1999) and a consensus of characteristic amino acids in specific positions of the homeodomain of known sequences (Pendleton et al., 1993; Ruddle et al., 1994a). Phylogenetic trees were constructed from partial homeodomain sequences using protein parsimony criteria implemented by PHYLIP 3.57c (J. Felsenstein, University of Washington, Seattle, 1995). The majority rule consensus tree presented was the outcome of 100 bootstrap sample analyses (PHYLIP 3.57c; Felsenstein, 1985). The following sequences (with their accession numbers in parentheses) were used in our similarity and phylogenetic analyses: nematode *Pristionchus pacificus* LIN-39 (AF052054); *Drosophila's lab* gene (P10105), *Ubx* (X76210) and *pb* (P31264); spider *Cupiennius salei* hox3 (AJ005643); sea star *Asterina minor Hox*-type homeobox fragments AM-1 (D86361), AM-3 (D86362), AM-5 (D86364), and AM-9 (D86367); sea urchin *Holopneustes purpurescens* (Hp) Hbox4 (U83419) and Hbox6 (U83420); sea urchin *Heliocidaris erythrogramma* (He) Hbox7 (U31564), Hbox9 (U31563), and Hbox10 (U31600); sea urchin *Paracentrotus lividus* (PI) Hbox11 (accession number not available); sea urchin *Tripneustes gratilla* (Tg) Hbox1 (M26370), Hbox3 (P10178), and Hbox4 (P10179); sea urchin *Strongylocentrotus purpuratus* (Sp) Hbox7 (AF042652), *Hox*1 (accession number not available), *Hox*2 (accession number not available), *Hox*4/5 (accession number not available), and *Hox*11/13a (accession number not available); acorn worm *Saccoglossus kowalevskii* Hox 9i (A44641) and SASHOXS8en (L1487); amphioxus *Brachiostoma floridaes* Hox2 (Z35143), Hox3 (X68045), Hox4 (Z45144), Hox 6/7 (L14880), Hox9 (Z35149), and Hox10 (Z35150); sea lamprey *Petromyzon marinus* Hox-6/7 (L14898) and Hox 9t (L14894); zebra fish *Danio rerio* Hoxd4 (Y14548); axolotl Hoxb-1 (P31357); *Xenopus* HOXB4 (P09070); chicken (c) Hoxb1 (P31259); mouse (m) Hoxa1 (M22115), Hoxa2 (M87801), Hoxb2 (M34004), Hoxd3 (X73573), Hoxc4 (X69019), Hoxb5 (M26283), Hoxa7 (M17192), Hoxa9 (AB005457), Hoxb9 (P20615), Hoxd9 (P28357), and Hoxc11 (P31312); and human (h) Hoxa-9 (AC004080) and Hoxc-9 (P31274).

Results

PCR amplifications with the degenerate primers *Hox* E and *Hox* F (Pendleton et al., 1993) were performed to isolate homeobox sequences from genomic DNA of the sea cucumber *H. glaberrima.* The expected 140-bp band was confirmed by gel electrophoresis. Six unique 82-bp homeobox fragments were identified from 38 clones analyzed. In addition, PCR amplifications with the degenerate primers 27-mer and 40-mer (Mito and Endo, 1997) were performed using a cDNA library of regenerating gastrointestinal tracts of *H. glaberrima* as template. Three unique 76-bp homeobox fragments were identified from 72 clones analyzed. The following *H. glaberrima* homeobox (HgHbox) sequences (with their corresponding accession numbers) were reported to the GenBank database: HgHbox1 (AF075751), HgHbox2 (AF075752), HgHbox3 (AF075755), HgHbox5 (AF075749), HgHbox9 (AF075750), HgHbox10 (AF075748), HgHBox11 (AF075753), HgHbox12 (AF075754), and HgHbox13 (AF075747).

The high conservation of the homeobox sequences, specifically the region encoding amino acids 21 to 47, has been used before to predict possible orthologous genes between *Hox* clusters (Murtha et al., 1991; Pendleton et al., 1993; Ruddle et al., 1994a; Mito and Endo, 1997; Morris et al., 1997). Table 1 shows a comparison at the amino acid level of partial homeodomain sequences of *H. glaberrima* with those of other species. HgHbox1 (2 clones), HgHbox2 (4 clones), and HgHbox3 (1 clone) showed the highest similarity to anterior group sequences. HgHbox1 has the highest degree of similarity to *Hox*1 sequences, while HgHbox3 has highest degree of similarity to *Hox*3 genes. HgHbox2 presents an interesting case, showing the highest similarity to cognate group 3 genes, but less than that of HgHbox3. HgHbox2 also shows similarity to cognate group 4 and 2 genes. We have placed it within the anterior group category and have named it HgHbox2, since HgHbox1 and HgHbox3 are better candidates for the holothurian cognates of anterior group 1 and 3 genes (see "Discussion"). HgHbox5 (1 clone) has the highest degree of amino acid similarity to the medial paralogy groups, *Hox* 5–7. Finally, HgHbox9 (1 clone), HgHbox10 (3 clones), HgHbox11 (1 clone), HgHbox12 (1 clone), and HgHbox13 (5 clones) show the highest degree of similarity to posterior paralogy group genes, the *Abd-B*-like genes, *Hox* 9–13. HgHbox9 is a *Hox*9-like sequence with the highest similarity to the sea urchin Hbox4 (Dolecki et al., 1986, 1988; Ruddle et al., 1994a; Morris et al., 1997) and the vertebrate Hoxa9 (Rubin et al., 1987) genes. HgHbox10 and HgHbox11 have the highest similarity to sea urchin posterior group genes Hbox10 (Ruddle et al., 1994a; Popodi et al., 1996; Morris et al., 1997; Martinez et al., 1999) and Hbox7 (Zhao, 1992; Ruddle et al., 1994a; Doblas et al., 1996; Popodi et al., 1996; Morris et al., 1997). In addition, we found two other sequences, HgHbox12 and HgHbox13, highly similar to each other and to echinoderm posterior genes, but considerably divergent from other *Hox* genes or homeobox-containing genes.

We performed a phylogenetic analysis using the protein parsimony criterion of the partial homeodomains (amino acids 21–27) from *H. glaberrima Hox*-like genes and those of other echinoderms, *Drosophila,* amphioxus, and mouse. Figure 1 shows an unrooted consensus tree in which *H. glaberrima* sequences are disperse. The tree shows four distinct branches representing four subfamilies of phylogenetically related genes: the anterior group has two branches, one with *Hox*1 sequences and the other with *Hox*2-*Hox*3 sequences; the medial group includes *Hox*4–8 sequences; and the posterior group, *Hox*9–13 sequences. The bootstrap percentage values (greater than 40%) are shown next to each branch. Bootstrap values supporting the main branches range from 43% to 100% except for the *Hox*2/3 branch, which had a low bootstrap value. The low value presented by the *Hox*2/3 branch may be related to the tendency of the echinoderm *Hox*2/3 sequences to form separate branches that do not correlate with their proposed position within the cluster. Furthermore, the conserved nature of the homeobox, together with the small number of character changes within the 27 amino acid span studied, might be responsible for the generally low bootstrap values obtained. Therefore, the consensus tree is a rough estimate of possible phylogenetic relationships among these genes (Felsenstein, 1985). However, this analysis sorts the nine *H. glaberrima Hox* sequences together with other known sequences into the assigned anterior, medial, and posterior groups shown on Table 1. Additional information is needed to allocate the sea cucumber sequences to definite cluster position, but our data indicate that there are three sequences in the anterior group, one sequence in the medial group, and five sequences in the posterior group genes within the putative holothuroid *Hox* cluster.

Discussion

We have employed the method of PCR amplification with degenerate primers to isolate homeodomain gene sequences from the sea cucumber *Holothuria glaberrima.* The sequences isolated are unique and show a high percentage of partial homeodomain similarity, indicating possible homology, with that of other echinoderms and metazoans. The homeodomain amino acid conservation allowed us to predict the distribution of the nine sea cucumber sequences within the anterior (HgHbox1, HgHbox2, and HgHbox3), medial (HgHbox5), and posterior (HgHbox9, HgHbox10, HgHbox11, HgHbox12, and HgHbox13) groups (as defined by Ruddle et al., 1994a). Our analysis is based on the procedure used by Murtha et al. (1991) and followed by several others (Pendleton et al., 1993; Ruddle et al., 1994a; Mito and Endo, 1997; Morris et al., 1997).

Evidence points out that in echinoderms (Popodi et al., 1996; Martinez et al., 1999) and, further, within most invertebrates studied to date (Caroll, 1995) there is a single *Hox* cluster. In the sea urchin *Hox* cluster, members from cognate groups 1, 2, and 3 have been documented (Di Bernardo et al., 1994; Popodi et al., 1996; Morris et al., 1997;

Table 1. Comparison of Partial Homeodomains of *Holothuria glaberrima* with Other Species*

Gene	Organism	Amino acid sequence	Similarity (%)	Source
		21 47		
Antennapedia	*Drosophila*	HFNRYLTRRRRIEIAHALCLTERQIKI		
Anterior group				
HgHbox1	Sea cucumber	---K----A--V---AM-E-N-T-V--		gDNA
SpHox1†	Sea urchin	---K----A--V---AS-Q-N-T-V	92	
AM-1†	Sea star	---K----A--V---AM-G-N-T-V	96	
lab	*Drosophila*	--------A------NT-Q-N-T-V--	81	
Hoxb1	Axolotl	---K----A--V---AT-E-N-T-V--	96	
cHoxb1	Chicken	---K----A--V---AT-E-N-T-V--	96	
Hoxa1	Mouse	---K----A--V---AS-E-N-T-V--	96	
HgHbox2	Sea cucumber	L-----H-S----M-ST-K-S------		gDNA
PlHbox11	Sea urchin	------C-P--V-M-KS-N--------	70	
AM-3†	Sea star	------G-P----M-GS-S------	76	
hox3	Spider	------C-P----M-NL-N-S------	78	
Hoxb3	Mouse	------C-P--V-M-NL-N-S------	74	
Hoxd4	Zebra fish	------------S--T-S-S------	78	
HOX-B4	*Xenopus*	-Y---------V----T-R-S------	66	
SpHox2†	Sea urchin	RL-H--C-P---Q--AY-E-S---V	56	
pb	*Drosophila*	---K--C-P------AS-D-----V-V	52	
Hoxa2	Mouse	---K--C-P--V---AL-D-----V-V	52	
HgHbox3†	Sea cucumber	------C-P--V-M-KS-N------		cDNA
PlHbox11	Sea urchin	------C-P--V-M-KS-N--------	100	
AM-3†	Sea star	------G-P----M-GS-S------	92	
pb	*Drosophila*	---K--C-P------AS-D-----V-V	78	
Hox3	Amphioxus	------C-P--V-M-AM-N--------	92	
Hoxd3	Mouse	------C-P--V-M-NL-N--------	92	
Medial group				
HgHbox5	Sea cucumber	--S---------------G--------		gDNA
HeHbox9	Sea urchin	------------------G--------	96	
SpHox4/5†	Sea urchin	------------------G------	96	
AM-5†	Sea star	------------------G------	96	
Hox6/7	Amphioxus	---K--------------G--------	92	
Hox6/7	Lamprey	------------------G--------	96	
Hoxa7	Mouse	--------------------------	92	
Posterior group				
HgHbox9	Sea cucumber	L--M----D--V---RL-N-----V--		gDNA
TgHbox-4	Sea urchin	L--M----D--L---RL-S-----V--	92	
HpHbox-4	Sea urchin	L--M----D--L---RL-T-----V--	92	
Hox-9i	Acorn worm	L--M----E--VD--RL-N-----V--	92	
Hoxa-9	Mouse	L--M----D--Y-V-RL-N-----V--	92	
Hoxa-9	Human	L--M----D--Y-V-RL-N-----V--	92	
HgHbox10	Sea cucumber	LY-M----D--SH-SR--S-----V--		gDNA
HeHbox10	Sea urchin	LY-M----D--SH-SR--S-----V--	100	
SpHox11/13a†	Sea urchin	LY-M----D--SH-SR--S-----V	100	
Hox-9t	Lamprey	L--M----D--Y-V-RV-S-----V--	77	
Hoxa-9	Mouse	L--M----D--Y-V-RV-N-----V--	74	
Hoxc-9	Human	L--M----D--Y-V-RV-N-----V--	74	

Table 1. Continued

Gene	Organism	Amino acid sequence	Similarity (%)	Source
HgHbox11†	Sea cucumber	QA-M----D--SKLSQ--D----LV		cDNA
HeHbox-7	Sea urchin	QA-M----D--SKLSQ--D-----V--	96	
SpHbox7	Sea urchin	TT-M----D--SKLSQ--D-----V--	88	
SASHoxs8en	Acorn worm	QQ-M----D--SRLSQ--N-----V--	80	
Hoxd-9	Mouse	L--M----D--Y-V-RI-N-----V--	60	
HgHbox-12†	Sea cucumber	QH-M----D--AKLSQT-S-----V		cDNA
HeHbox-7	Sea urchin	QA-M----D--SKLSQ--D-----V--	84	
SpHbox-7	Sea urchin	TT-M----D--SKLSQ--D-----V--	80	
Sashoxs8en	Acorn worm	QQ-M----D--SRLSQ--N--------	76	
Hox-9	Amphioxus	LY-M----E--Y--SQHVN-----V--	64	
Hoxc-11	Mouse	L--M----E--L--SKTIN--D--V--	60	
HgHbox-13	Sea cucumber	KAQQ----D--ARLSQS-S-S---V--		gDNA
AM-9†	Sea star	QAHQ----D--TRL-QS-S-----V	80	
SpHbox-7	Sea urchin	TT-M----D--SKLSQ--D-----V--	68	
lin-39	Nematode	---K----K-----S-S-M-S---V--	60	
SASHOXS8en	Acorn worm	QQHM----D--SRLSQ--N--------	60	
Hoxb-9	Mouse	L--M----D--H-V-RL-N-S---V--	56	

*Amino acids 21 to 47 of the homeodomains of *H. glaberrima Hox*-like sequences (HgHbox) are aligned with the most similar of other echinoderms and animal groups (†with the exception of HgHbox3, HgHbox11, HgHbox12; *Strongylocentrotus purpuratus* Sp*Hox*1, Sp*Hox*2, Sp*Hox*4/5, and Sp*Hox*11/13a; and *Asterina minor* AM-1, Am-6, and AM-9, in which available data only contain amino acids 21–45 of the homeodomain). Amino acids identical to those of the *Drosophila's Antennapedia* (*Antp*) are indicated by a dash. Percentage similarity is in relation to the *H. glaberrima* sequences. The HgHbox sequences are listed under group headings (anterior, medial, and posterior) in accordance with the putative echinoderm clusters reported by Popodi et al. (1996), Mito and Endo (1997), and Martinez et al. (1999). For sequence information refer to cloning and sequencing section in "Materials and Methods".

Martinez et al., 1999). In the sea stars, AM-1 and AM-3, partial homeodomain sequences corresponding to cognate groups 1 and 3, respectively, have been reported (Mito and Endo, 1997). We are documenting the presence of partial homeodomain sequences of a *Hox*1-like sequence, HgHbox1, and a Hox3-like sequence, HgHbox3, in sea cucumbers. HgHbox1 has 92% amino acids sequence similarity to Sp*Hox*1, and 96% to AM-1 and other *Hox*1 genes. HgHbox3 has 100% amino acid sequence similarity to sea urchin Hbox11, and it is highly similar to AM-3 and *Hox*3 genes from other groups. Such high similarity provides strong support for the homology of these genes.

In addition, we report here a *Hox*2-like sequence, HgHbox2. The position of HgHbox2 sequence is more difficult to explain, being highly similar to *Hox*3-like sequences from other organisms, but analogous situations have been recorded in the literature before (Pendleton et al., 1993). A possible explanation is that PCR errors or polymorphism may lead to variation in gene products. However, our results suggest that this is unlikely since several HgHbox2 clones were isolated and all were shown to be identical, which should not be the case if the sequence were caused by PCR mistakes in the amplification of HgHbox3. Furthermore, although HgHbox3 came from an expression library and HgHbox2 is from genomic DNA, the differences between these sequences cannot be explained by intron splicing rearrangement even though there have been reports of introns disrupting the homeobox in *Drosophila*'s anterior genes. (Ruddle et al., 1994a). A plausible explanation might be that HgHbox2 is a *Hox*3-derived sequence (Zhang and Nei, 1996; Brooke et al., 1998; Purugganan, 1998), the outcome of a recent tandem duplication of the *Hox*3 gene (Kappen and Ruddle, 1993).

Therefore, HgHbox2 might represent a divergent *Hox*2 or a *Hox*4 sequence within the holothuroid cluster. We have chosen to place it as the putative HgHbox2 because sequence analysis places it among the anterior group members and HgHbox5 has stronger similarity to the echinoid *Hox*4/5 gene previously reported (Ruddle et al., 1994a; Martinez et al., 1999). In addition, in echinoids *Hox*2 has also been found to be less similar to the *Drosophila* and vertebrate *Hox*2 genes (~72% and 69% similarity, respectively)

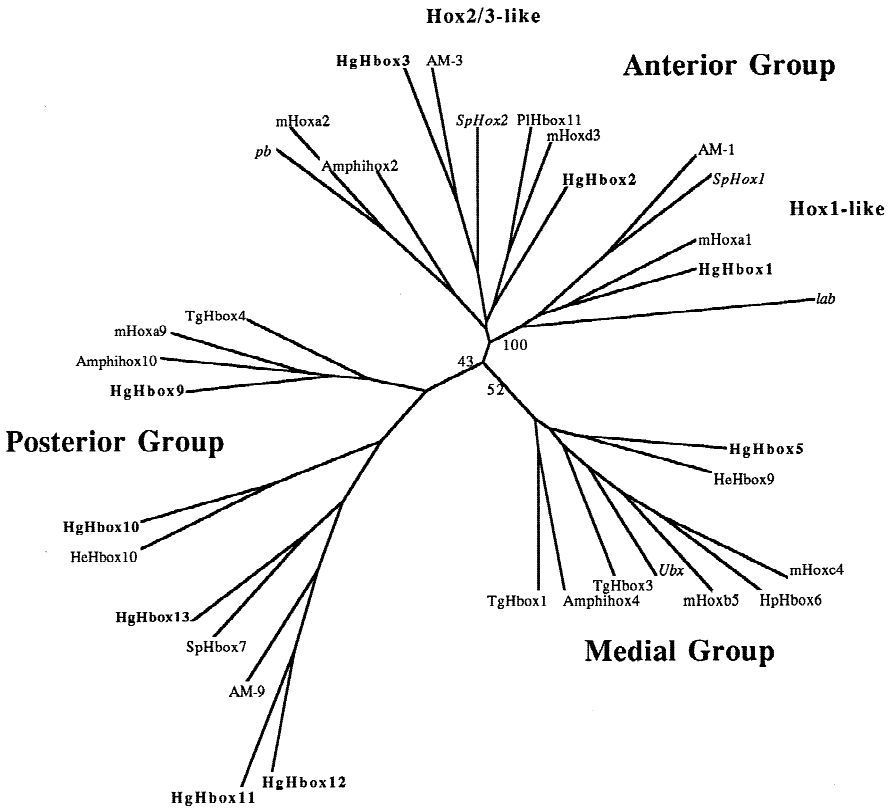

Figure 1. Phylogenetic relation of *Holothuria glaberrima* partial homeodomain sequences (bold) to other echinoderm, *Drosophila,* amphioxus, and mouse sequences. This is a majority-rule consensus tree generated by bootstrap estimate of phylogenies evaluated by the protein parsimony criterion using PHYLIP 3.57c (J. Felsenstein, 1995). The tree is unrooted and shows four subfamily branches representing the anterior, medial, and posterior groups. The numbers indicate bootstrap values (over 40%) of main branches. m indicates mouse; Amphihox, amphioxus (*Brachiostoma floridae*) *Hox* gene; AM, sea star *Asterina minor Hox* sequence; Hbox, sea urchins *Hox* sequence (species: PI, *Paracentrotus lividus;* Sp, *Strongylocentrotus purpuratus;* Hp, *Holopneustes purpurescens;* He, *Helicidaris erythrogramma;* Tg, *Tripneustes gratilla*) and *Drosophila's lab, labial; pb, proboscipedia;* and *Ubx, Ultrabithorax.*

and also falls, in our phylogenetic analysis, with the *Hox*3 cognates. Finally, we cannot rule out the possibility that the HgHbox2 sequence is not a homeobox gene but rather some other homeodomain not part of the *Hox* cluster. Nevertheless, HgHbox2 does not show any similarity with other known non-*Hox*-type genes (data not shown). If future chromosome walking confirms the location of this *Hox*2 gene within the holothurian cluster, it is tempting to speculate that the large alterations in the sequence of *Hox*2 genes both in echinoids an holothurians might be responsible for some of the morphological characteristics unique to these echinoderm groups.

Only one sequence was isolated from the medial group, but we believe that there are other medial group genes within the sea cucumbers. A total of four different medial group genes have been reported in echinoderms (Popodi et al., 1996; Mito and Endo, 1997; Martinez et al., 1999). Thus, it seems that the medial group genes in our PCR surveys are underrepresented. The cluster gene recovery rate by PCR surveys has been estimated as more than 85% (Pendleton et al., 1993), but this is not always the case. Several of the already-identified cluster sequences were missing in analogous echinoderm experiments (Ruddle et al., 1994a; Popodi et al., 1996; Mito and Endo, 1997; Morris et al., 1997).

In addition, we report the isolation of five members (HgHbox9, HgHbox10, HgHbox11, HgHbox12, and HgHbox13) of the *Abd-B*-like posterior group genes. As expected, these sequences are highly similar among themselves and considerably divergent from other members of the putative cluster. In addition, they have a high percentage of similarity (>80%) to echinoderm posterior sequences, Hbox4, Hbox7, Hbox10, and AM-9, indicating possible homology. These relations, as well as a relation to mouse *Hoxa*9 and amphixus *Hox*9 genes, are shown in our phylogenetic analysis. This is the first documentation of evidence sustaining the presence of five posterior sequences in echinoderms. In fact, in their recent paper showing the organization of the echinoid *Hox* cluster, Martinez and colleagues (1999) found only three posterior groups. Our finding strengthens the work of Popodi et al. (1996), who proposed that the expansion of genes in the *Abd-B* group occurred before the divergence of vertebrates from other deuterostomes. Members of the class Holothuroidea, together with echinoids, are thought to be among the most evolutionarily derived members of the Echinodermata. They maintain a pentaradial symmetry, but also posses a secondarily derived bilateral symmetry. Recent studies on echinoderm homeobox genes, *distalless, engrailed,* and *or-*

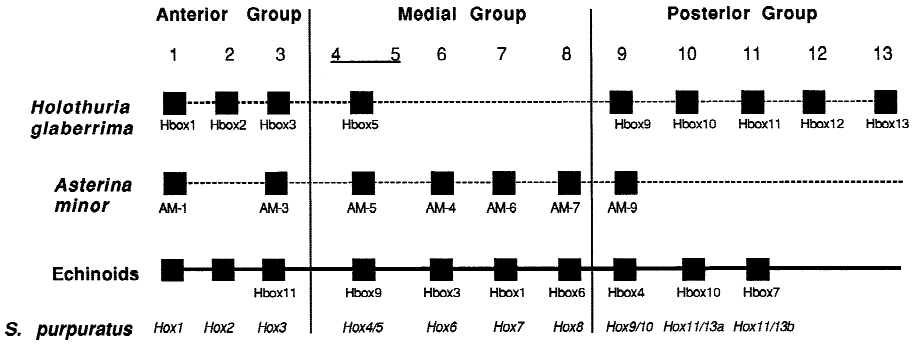

Figure 2. Putative *Holothuria glaberrima Hox* cluster in relation to the ones presented for the echinoids (Popodi et al., 1996; Martinez et al., 1999) and the sea star *Asterina minor* (Mito and Endo, 1997) echinoderms. The clusters are subdivided in the anterior, medial, and posterior group genes. The numbers at the top represent the paralog group position of vertebrate *Hox* clusters. Dashes indicate the putative nature of the cluster organization in *H. glaberrima* and *A. minor.* The organization of genes in echinoids is shown on top from the original characterization of the gene sequences (Popodi et al., 1996). Below is the *Strongylocentrotus purpuratus Hox* cluster, the only one in which paralog groups have been experimentally confirmed by genomic walking (Martinez et al., 1999).

thodenticle, showed that their role in echinoderms differs from their roles in other animal groups and, moreover, even between echinoderms classes (Lowe and Wray, 1997; Davidson, 1997). It might be plausible, then, that some of the characteristics particular to the class Holothuroidea are dependent on the posterior *Hox* sequences that have been characterized here, although this cannot be known without further experiments.

This analysis provides an insight of the *Hox*-type genes within the holothuroids and allows us to propose a possible cluster among these genes (see Figure 2). As suggested previously for the echinoid Hox cluster (Popodi et al., 1996; Martinez et al., 1999), the proposed holothurian cluster is more similar to those of chordates than to the ones reported from other invertebrates, such as *Drosophila.* Verification of a cluster arrangement must be completed, however. Further information, such as the complete homeobox gene and protein sequences, gene locus, and expression patterns, is required to establish the exact nature of these genes.

Acknowledgments

We thank Dr. Michael Rubin for his support and experimental advice, and Ms. Lucybeth Nieves for her excellent technical assistance. We also thank Mr. Humberto Ortiz-Zuazaga, Dr. Owen McMillan, and Dr. Stephen Rehner for their help in the phylogenetic tree analysis. This work was funded by EPSCoR, NIGMS-MBRS, RCMI, MSIP, and the UPR.

References

Akam, M. (1989). *Hox* and Hom: homologous gene clusters in insects and vertebrates. *Cell* 57:347–349.

Amores, A., Force, A., Yan, Y., Joly, L., Amemiya, C., Fritz, A., Ho, R. K., Lagenland, J., Prince, V., Wang, Y., Westerfield, M., Ekker, M., and Postlethwait, J.H. (1998). Zebra *hox* clusters and vertebrate genome evolution. *Science* 282:1711–1714.

Aviv, H., and Leder, P. (1972). Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proc Natl Acad Sci USA* 69:1408–1412.

Balavoine, G. (1998). Are platyhelminthes coelomate without a coetom? An argument based on evolution of Hox genes. *Am Zool* 38:843–858.

Bayascas, J.R., Castillo, E., Muñoz-Mármol, A.M., and Saló, E. (1997). Planarian Hox genes: novel patterns of expression during regeneration. *Development* 124:141–148.

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

Carroll, S.B. (1995). Homeotic genes and the evolution of arthropods and chordates. *Nature* 376:479–485.

Cartwright, P., Dick, M., and Buss, L.M. (1993). Hom/*Hox* type homeoboxes in chelicerates *Limulus polyphenus. Mol Phylogenet Evol* 2:85–192.

Chomczynski, P., and Sacchi, N. (1987). Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 162:156–159.

Davidson, E.H. (1998). Insights from the echinoderms. *Nature* 389:679–680.

Degnan, B.M., Degnan, S.M., Guisti, A., and Morse, D.E. (1995). A Hom/Hox homeobox gene in sponges. *Gene* 155:175–177.

Di Bernardo, M., Russo, R., Olivieri, P., Melfi, R., and Spinelli, G. (1994). Expression of homeobox-containing genes in the sea urchin (*Paracentrotus lividus*) embryo. *Genetica* 94:141–150.

Dobias, S.L., Zhao, A.Z., Tan, H., Bell, J.R., and Maxson, R. (1996). *SpHbox7,* a new *Abd-B* class homeobox gene from the sea urchin *Strongylocentrotus purpuratus:* insights into the evolution of *Hox* gene expression and fuction. *Dev Dynam* 207:450–460.

Dolecki, G.J., Wannakrairoj, S., Lum, R., Wang, G., Riley, H.D., Carlos, R., Wang, A., and Humphreys, T. (1986). Stage-specific expression of a homeo box-containing gene in the non-segmented sea urchin embryo. *EMBO J* 5:925–930.

Dolecki, G.J., Wang, G., and Humphreys, T. (1988). Stage- and tissue-specific expression of two homeobox genes in the sea urchin embryos and adults. *Nucleic Acids Res* 16:11543–115440.

Finnerty, J.R., and Martindale, M.Q. (1997). Homeobox in sea anemones (Cnidaria; Anthozoa): a PCR-based survey on *Nematostella vectensis* and *Metridium senile. Biol Bull* 193:62–76.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783–791.

García-Arrarás, J.E., Estrada-Rodgers, L., Santiago, R., Torres, I.I., Díaz-Miranda, L., Torres-Avillán, I. (1998). Cellular mechanisms of intestine regeneration in the sea cucumber, *Holothuria glaberrima* Selenka (Holothuroidea: Echinodermata). *J Exp Zool* 281: 288–304.

Gehring, W.J. (1992). The homeobox perspective. *Trends Biochem Sci* 17:277–280.

Holland, P.W.H. (1998). Major transitions in animal evolution: a developmental genetic perspective. *Am Zool* 38:829–842.

Holland, P.W.H., and García-Fernàndez, J. (1996). *Hox* genes and the chordate evolution. *Dev Biol* 173:382–395.

Kappen, C., and Ruddle, F.H. (1993). Evolution of regulatory gene family; HOM/*Hox* genes. *Curr Opin Genet Dev* 3:931–938.

Lowe, C.J., and Wray, G.A. (1997). Radical alterations in the roles of homeobox genes during echinoderm evolution. *Nature* 389: 718–721.

Martinez, D.E., Bridge, D., Masuda-Nakagama, L.M., and Cartwright, P. (1988). Cnidarian homeoboxes and the zootype. *Nature* 393:748–749.

Martinez, P., Rast, J.P., Arenas-Mena, C., and Davidson, E.H. (1999). Organization of an echinoderm *Hox* gene cluster. *Proc Natl Acad Sci USA* 96:1469–1474.

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

Mito, T., and Endo, K. (1997). A PCR survey of *Hox* genes in the sea star, *Asterina minor. Mol Phyl Evol* 8:218–224.

Morris, V.B., Brammall, J., Byrne, M., and Frommer, M. (1997). *Hox*-type and non-*Hox* homeobox gene sequences in the genomic DNA of the sea urchin *Holopneustes purpurescens. Gene* 201:107–110.

Murtha, M.T., Leckman, J.F., and Ruddle, F.H. (1991). Detection of homeobox genes in development and evolution. *Proc Natl Acad Sci USA* 88:10711–10715.

Pendleton, J.W., Nagai, B.K., Murtha, M.T., and Ruddle, F.H. (1993). Expansion of the *Hox* gene family and the evolution of chordates. *Proc Natl Acad Sci USA* 90:6300–6304.

Popodi, E., Kissinger, J.C., Andrews, M.E., and Raff, R.A. (1996). Sea urchins *Hox* genes: insights into ancestral *Hox* cluster. *Mol Biol Evol* 13:1078–1086.

Purugganan, M.D. (1998). The molecular evolution of development. *Bioessays* 20:700–711.

Rubin, M.R., King, W., Toth, L.E., Sawczuk, I.S., Levine, M.S., D'Eustachio, P., and Nguyen-Hu, M.C. (1987). Murine Hox-1.7 homeo-box gene: cloning, chromosomal location, and expression. *Mol Cell Biol* 7:3836–3841.

Ruddle, F.H., Bentley, K.L., Murtha, M.T., and Risch, N. (1994a). Gene loss and gain in the evolution of the vertebrates. *Development Suppl* 155–161.

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

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74: 5463–5467.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

Scott, M.P. (1993). A rational nomenclature for vertebrate homeobox (*HOX*). *Nucleic Acids Res* 21:1687–1688.

Tarabykin, V.S., Lukyanov, V.K., and Lukyanov, S.A. (1995). Detection of planarian Antennapedia-like homeobox genes expressed during regeneration. *Gene* 158:197–202.

Wang, G.V.L., Dolecki, G.J., Carlos, R., and Humphreys, T. (1990). Characterization and expression of two sea urchin homeobox gene sequences. *Dev Gen* 11:77–87.

Washabaugh, C.H., Wallace, J.L., Del Rio-Tsonis, K., and Tsonis, P. (1995). Cloning of homeobox sequences expressed in intact and the regenerating newt eye. *Gene* 158:301–302.

Zhao, A.Z. (1992). Temporal regulation of the L1 late H2B histone gene during sea urchin embryonic development by an antennapedia class homeoprotein. Ph.D. dissertation, University of Southern California, Department of Biology, Los Angeles, Calif.

Zhang, J.Z., and Nei, M. (1996). Evolution of the *Antennapedia*-class homeobox genes. *Genetics* 142:295–303.