J Mol Evol (2003) 56:509–516
DOI: 10.1007/s00239-002-2421-2

JOURNAL OF MOLECULAR EVOLUTION

Letter to the Editor

Gene Fishing: The Use of a Simple Protocol to Isolate Multiple Homeodomain Classes from Diverse Invertebrate Taxa

Shannon E. Lee, Ruth D. Gates, David K. Jacobs

Department of Organismic Biology, Ecology, and Evolution, University of California, Los Angeles, 621 Charles E. Young Drive South, Los Angeles, CA 90095-1606, USA

Received: 14 April 2002 / Accepted: 3 December 2002

Abstract. Comparison of relevant gene sequence and functional data is central to understanding the evolution of metazoan development. The conservation of portions of regulatory genes, such as homeoboxes, allows for the design of PCR-based sequence isolation and amplification strategies. Here we describe a simple protocol that uses a degenerate primer pair to isolate a variety of homeobox-containing genes from diverse metazoan taxa. In a non-exhaustive survey, we have isolated 28 gene sequence fragments from 15 taxa, representing eight invertebrate phyla (Mollusca, Echiura, Annelida, Platyhelminth, Acoela, Ctenophora, Cnidaria, and Porifera). Based on BLAST and parsimony analyses, these gene fragments affiliate with several gene groups (PAIRED-like, HOX, and ParaHOX) and several single genes, including pancreas/duodenum homeoboxes (*Pdx*), empty spiracles (*ems*/*Emx*), gastrulation brain homeoboxes (*Gbx*), hematopoietically expressed homeoboxes (*HEX*), brain specific homeobox (*bsh*/*BarH1*/*BarH2*), NK-1 (*NK-1*/*s59*/*slouch*), and ladybird (*Lbl*/*Lbe*/*Lbx*). In several cases, these fragments represent the first reported orthologue for the phylum or superphyletic group (i.e., Lophotrochozoa).

Key words: Lophotrochozoa — Homeobox genes — PCR — HOX — ParaHOX — PAIRED — Invertebrate — Metazoa

Comparative studies of developmental genes provide critical information about the morphological evolution of the metazoa. The sequences of many of these genes contain motifs that are highly conserved across taxa, a characteristic that allows investigators to design PCR-based strategies to recover gene fragments from unexplored organisms. For example, the often referred to "Ruddle" primers, and other similar protocols, have been used successfully to isolate small fragments of Hox, and other related, homeobox-containing genes (Murtha et al. 1991; Schierwater et al. 1991; Bartels et al. 1993; Cartwright et al. 1993; Pendleton et al. 1993; Dick and Buss 1994; Snow and Buss 1994; Finnerty et al. 1996; de Rosa et al. 1998; Mouchel-Viehl et al. 1998; Abzhanov and Kaufman 2000; Kim et al. 2000; Bastianello et al. 2002).

Although the amount of information regarding developmental gene sequence and function continues to grow, these data are primarily generated for classic developmental systems such as mouse, fly, and nematode, and many phyla of evolutionary interest remain unexplored. Here we describe a simple PCR-based technique that isolates a diversity of homeod-

Correspondence to: Shannon E. Lee; *email:* shannond@ucla.edu

Table 2. Taxonomic and gene affiliation information for the novel gene fragments[a]

Phylum	Species name	Gene affiliation(s)[b]
Mollusca	*Crassostrea gigas*	PAIRED-type (I)
	Nutricola sp. (cDNA)	*empty spiracles/Emx*
	Kelletia kelletii	*ind/Gsh*, *pdx*
	Bulla bulla	HOX
	Loligo opalescens	GBX2, *ladybird/Lbx*
Echiura	*Urechis caupo*	PAIRED-type (I), HOX, *empty spiracles/EMX*
	Urechis caupo (cDNA)	GBX2
Annelida	*Lumbricus* sp.	HOX, GBX2
Platyhelminthes	*Notoplana acticola*	PAIRED-type, HOX
	Stylochus tripartitus	PAIRED-type
	Himasthla rhigedana	*NK-1/s59*
Acoela	*Neochildia fusca*	HOX
Ctenophora	*Beroe* sp.	PAIRED-type, HEX/BAR
Cnidaria	*Aurelia aurita*	HOX
Porifera	*Acarnus* sp.	PAIRED-type
	Spongilla sp.	PAIRED-type (I)

[a] Summary of gene isolation data. Results are listed according to taxonomy and organized by phylum and species name. In addition, the gene affiliation of the isolated genes is indicated. Except where indicated as (cDNA), genomic DNA was used as the template in the PCR amplifications.
[b] I, intron present.

omain-containing gene fragments (PAIRED-like, HOX, ParaHOX, and others) from a broad range of metazoans. This protocol differs from the aforementioned techniques in that it isolates a greater variety of homeodomain containing genes (i.e., well beyond the HOX subset) and produces a slightly larger fragment.

Our protocol, and the "Ruddle"-type techniques, generate fragments that are less than 100 base pairs in length and likely too small to be directly useful as a probes in functional assays. However, these gene fragments allow for (1) a first level examination of gene sequences from understudied taxa, (2) a phylogenetic comparison of sequences across diverse taxa, and (3) the subsequent design of species and gene specific PCR primers for use in gene extension protocols. Additionally, when generated using tissue- or stage-specific cDNA templates and rt-PCR techniques, these sequences provide preliminary information on developmental timing and regional specificity. As such, these techniques have broad utility.

Our protocol employs two degenerate primers that correspond to the amino acid translations QLEELE (5′ cagctsgargarctggag 3′) and QNRRA (5′ gcbckncgrttytgraacc 3′) to amplify 96 bp of novel homeodomain sequence (novel = excluding primers). We have used these primers to obtain multiple homeodomain gene fragments from cDNA and genomic DNA templates extracted from five molluscs, one echiuran, one annelid, three flatworms, one acoel flatworm, one ctenophore, one cnidarian, and two sponges (see web-based Table 1 [http://www.link.springer-ny.com/link/service/journals/00239/index.htm] for taxonomic position, collection site, and template type.

RNA was extracted with Trizol (GibcoBRL) and reverse-transcribed to cDNA using Superscript II (GibcoBRL) according to the manufacturer's instructions, and genomic DNAs were purified using a standard phenol–chloroform extraction. These templates were used in 50-µl PCR reactions containing 1 × PCR buffer, 2 m*M* $MgCl_2$, 1 U AmpliTaq Gold DNA polymerase (Applied BioSystems), 0.2 µ*M* dNTPs (Amersham Pharmacia Biotech), and a 0.5 µ*M* concentration of each primer. The AmpliTaq Gold DNA polymerase was activated for 12 min at 94°C and the gene fragments amplified with 30 cycles of 1 min at 94°C, 1 min at 53°C, and 2 min at 72°C. A second round of PCR was conducted on 1 µl of the primary PCR amplification, including the negative control, using the cycling conditions described above. PCR products were cloned using the TOPO TA cloning kit (version K; Invitrogen) and the plasmid DNAs were isolated using the Qiagen Plasmid Minikit (Qiagen) according to the manufacturer's instructions. These clones were manually sequenced to screen for homeobox fragments and the homeobox-positive clones resequenced with an automated sequencer. Each homeobox-positive sequence was listed by genus and sequentially assigned a number (e.g., Stylochus-1). Sequences generated from cDNA templates were further designated with a "C" (e.g., Urechis-1C).

We have employed this protocol on eight phyla and amplified a total of 28 gene fragments belonging to nine gene groups (Table 2). Most isolated fragments were the expected 96 bp in length, however, three PCR products, amplified from genomic DNA templates, contained short (< 300 nucleic acids) introns (Table 2). Figure 1 aligns the deduced amino

Gene Name	Taxa	Amino Acid Sequence Alignment	Closest Blast	
Antp_Dm[X03791]	ARTHROPODA	KEFHFNRYLTRRRRIEIAHALCLTERQIKIWF	Antp_Fly	[X03791]
Urechis-1	ECHIURAN	-A-ERAP-PDVFA-E-L-LRIG-S-SRVQV--	AL_Fly	[Q06453]
Crassostrea-1	MOLLUSC	-A-AETH-PDIYT-E---MKTD---ARVQV--	PMXB_Mouse	[O35690]
Beroë-4	CTENOPHORE	-V-GRTH-PDLVV-E-L-GRIG-S-SCVQ---	PMX2_Human	[Q99811]
Stylochus-1	FLATWORM	-A-QEAH-PDVYH-EVLSIKTD-P-DR-QV--	VSX1_Fish	[Q90277]
Notoplana-2	FLATWORM	-A-QEAH-PDVYQ-EVLSIKTD-P-DR-QV--	VSX1_Fish	[Q90277]
Acarnus-1	SPONGE	AA-GKTH-PDVFM-EDL-MRIN---ARVQV--	DRG11_Rat	[Q62798]
Spongilla-1	SPONGE	AA-SKTH-PDVFM-EDL-LRIN---ARVQ???	DRG11_Rat	[Q62798]
Kelletia-5	MOLLUSC	------K----A------A--G-N-T-V----	HOXA1_Frog	[Q08821]
Notoplana-1	FLATWORM	---Y--K----A------N--G-N-T-V----	HOXD1_Frog	[Q08820]
Neochildea-1	ACOEL	------K----A--V---AS-Q-N-T-V----	HOXA1_Human	[P49639]
Bulla-1	MOLLUSC	------K--C-P------AS-D-----V-V--	PB_Fly	[P31264]
Lumbricus-2	ANNELID	???------C-P----M-AM-K--------??	HOX3_Cephal	[P50901]
Urechis-2	ECHIURAN	----------C-P----M-AL-N----------	HOX3_Cephal	[P50901]
Aurelia-1	CNIDARIAN	------HF---E--LQL-TN-N----------	ABDA_Beetle	[Q07961]
Kelletia-1	MOLLUSC	------K-IS-P----L-AM-S----H-----	HTRA2_Leech	[P17138]
Kelletia-6	MOLLUSC	------K-IS-PG---L-AM-S----H-----	HTRA2_Leech	[P17138]
Kelletia-9	MOLLUSC	------K-IS-P----L-AM-S---PH-----	HTRA2_Leech	[P17138]
Urechis-3	ECHIURAN	T--GR-Q--VG-E-SKL-QR-Q-S-T-V-V--	EMS_Fly	[P18488]
Nutricola-1C	MOLLUSC	R--QR-Q--VGKE-S-L-KR-R-S-T-V-V--	NK-3_Fly	[P22809]
Loligo-1	MOLLUSC	----SKK--SLTE-SQ---N-K-S-V-V----	GBX2_Mouse	[P48031]
Urechis-1C	ECHIURAN	----SKK--SLTE-SQ---Q-R-S-V-V----	GBX2_Mouse	[P48031]
Lumbricus-1	ANNELID	----SKK--SLTE-SH---Q-K-S-V-V---?	GBX2_Mouse	[P48031]
Beroë-1	CTENOPHORE	RK-RTQK--SVPE-LDE-TT---S-T-V-T--	HEX_Mouse	[P43120]
Beroë-2	CTENOPHORE	RK-RTQK--SVPE-LDV-TT---S-T-V-T--	HEX_Mouse	[P43120]
Himasthla-2	FLATWORM	NK-KTT---SVCE-LNL-PS-N---T-V----	NK-1_Fly	[P22807]
Loligo-2	MOLLUSC	-R-LYQK---PAD-D---LS-G--NA-VIT--	LBX1_Human	[P52954]
Kelletia-3	MOLLUSC	R--AA-M--S-L------TL-N-S-K-V----	GSH1_Mouse	[P31315]

Fig. 1. Putative amino acid translations for the gene sequence fragments aligned against the *Drosophila Antp* gene. *Dashes* indicate identical amino acid residues compared to the alignment standard. Closest BLAST results are listed to the right of the translations and indicate the top BLAST result for each translated new sequence using the SwisProt database, basic BLAST, and amino acid translations of the gene fragments as the search template. Accession numbers are given in *brackets* to the right of the gene name and the taxonomic affiliation.

acid translations for our gene fragments against the commonly used hox standard, *antp,* translated from the *Drosophila* sequence. To assign the gene fragments an affiliation, each sequence was compared to the SwissProt database using the standard protein-to-protein Basic Local Alignment Search Tool, BLAST (Blastp; http://www.ncbi.nlm.nih.gov/BLAST/ [Altschul et al. 1990]). The corresponding top BLAST result for each new gene is also listed in Fig. 1. Due to the dynamic nature of the international databases, more accurate affiliations may arise as new data become available.

To additionally examine the relationships of our gene fragments to other published gene sequences, we conducted parsimony analyses on a combined amino and nucleic acid database (Agosti et al. 1996; Jacobs et al. 1998; Schubert et al. 2000). To accomplish this, we initially compared our fragments with aligned portions of all *Drosophila* homeobox-containing genes (spanning the amino acid translation PKRKQ to QEKVGP of the *aristaless* gene [accession No. L08401]) using heuristic searches (third-base pair positions removed for the nucleic acid section of the combined database; PAUP version 4.0b10). Based on these analyses, we excluded the most basal genes or gene groups for whom we have no novel orthologues (i.e., iroquois, extradenticle, and sineoculis) and inferred appropriate outgroups (Lim3, zf1, IPOU) for our later analyses. We then added several vertebrate examples of homeobox-containing genes that are not well represented in the *Drosophila* genome (i.e., several PAIRED-like, Hex, Gbx, and IPF-related genes) and included any genes similar to our novel fragments (based on the BLAST surveys) isolated from the same or related phyla.

After third-base pair positions were removed from the nucleic acid section of the combined data set, the matrix was analyzed as follows. First, a heuristic search was conducted using only the *Drosophila* and vertebrate genes, with our gene fragments excluded (200 repetitions: Tree Bissection/Reconnection swapping algorithm). Second, the characters were reweighted to reflect the topology of this search (reweight characters function, base weights = 1000, maximum value/ best fit is more than one tree, and reweighted according to the rescaled consistency index, RC). This weighting strategy attempts to address the issue of the limited data and consequent resolution based on the short sequences recovered from the PCR protocol. Finally, our gene fragments, and the closely related orthologues, were added to the reweighted data set and subjected to a heuristic search (200 repetitions) and a jacknife analysis (50% deletion, emulate 'Jac' resampling, 100 repetitions, re-

taining groups with frequencies > 50% [Farris et al. 1996]). The resulting strict consensus tree, from the heuristic search, with jacknife values overlain is shown in the web-based Fig. 2.

The BLAST results (Fig. 1) and parsimony analyses (Fig. 2;) reveal that some of our gene fragments are similar to members of the PAIRED-like, HOX, and ParaHOX gene groups, and others are closely related to the *pancreas/duodenum homeoboxes (Pdx)*, *empty spiracles (ems/Emx)*, *gastrulation brain homeoboxes (Gbx)*, *hematopoietically expressed homeoboxes (HEX)* and/or *brain specific homeobox (bsh/BarH1/BarH2)*, *NK-1 (NK-1/s59/slouch)*, and *ladybird (Lbl/Lbe/Lbx)* genes. Throughout the text, our gene fragments are discussed in the context of the most closely related *Drosophila* gene (e.g., NK-1), however, where the gene is not represented in the *Drosophila* genome, we refer to the most closely related vertebrate orthologue (e.g., *Pdx*, *Gbx*).

PAIRED-like Genes. Seven of our gene fragments are PAIRED-like genes. To our knowledge, the echiuran sequence represents the first reported PAIRED-like gene for this phylum, and the molluscan, platyhelminth, and poriferan sequences appear to be novel types within each of these phyla. PAIRED-like genes are characterized by two conserved motifs, a homeodomain and a PAIRED domain. The homeodomain of this gene group has a distinctive "YPD" amino acid signature. Our gene fragments contain this signature, and according to the BLAST analysis, they are most closely related to *reversed polarity (repo)*, *Paired mesoderm homeobox (Pmxb/Phox2b/Pmx2)*, *Visual system homeobox* (*Vsx)*, and *Drg11*.

BLAST and parsimony analyses document the similarity between the echiuran gene Urechis-1 (AY187691) and the *Drosophila* gene *repo*. In *Drosophila,* the *repo* gene is expressed in most glial cells in both the CNS and the peripheral nervous system of embryonic flies (Xiong et al. 1994). A detailed examination of amino acid translations in the PAIRED-like gene group reveals that Urechis-1 also has signatures similar to those of the *Ptx* and *Mix* PAIRED-like genes (Galliot et al. 1999), suggesting possible alternative assignment.

BLAST results for the molluscan and ctenophore fragments (Crassostrea-1 [AY187692] and Beroe-4 [AY187693], respectively) document similarity to the vertebrate *Pmxb/Phox2b/Pmx2* genes. These genes play a role in the specification of neurons in both the central and the peripheral nervous systems (see Pattyn et al. 1997), but their role outside vertebrate model systems has not yet been explored. The Crassostrea-1 sequence, however, has signatures consistent with *al* and *Arix* genes (Gaillot et al. 1999), and in the parsimony analysis, this gene groups with *vsx* and *cart* genes. Thus, precise assignment is ambiguous. To our knowledge, this is the first report of a gene of this general type from the phylum Mollusca. A gene very similar to Beroe-4 has been reported from the congener *Beroe ovata* (*Ctenopaired* [U21206]). This gene is identical in amino acid translation to our ctenophore fragment, but only 85% similar in nucleic acid sequence (Finnerty et al. 1996), and in our parsimony analysis, the two genes cluster as sister sequences to the *cart* gene group member.

The two platyhelminth PAIRED-like fragments, Stylochus-1 (AY187694) and Notoplana-2 (AY187695), are most closely affiliated with the vertebrate *vsx* genes based both on the BLAST searches and on the parsimony analysis. Although there are a number of published PAIRED-like flatworm genes (i.e., Smox-3 [M85303], DTPRD1 [Xl0299], and EGHBX4 [Oliver *et al.* 1992]), they are not as similar to *vsx* genes as Stylochus-1 and Notoplana-1. The human counterpart of this gene is expressed in embryonic craniofacial and adult ocular tissues (Semina et al. 2000). The *C. elegans* gene, *ceh-10*, may be closely related to *vsx*, but we know of no other reports of *vsx* orthologues from invertebrates. It will be interesting to see if these flatworm orthologues are involved in the patterning of photoreceptors or anterior neurogenic elements within these taxa.

Our two sponge PAIRED-like fragments, Acarnus-1 (AY187696) and Spongilla-1 (AY187697), differ from other published PAIRED-like sponge genes (i.e., PAX-37 [D84254] and PAX 2/5/8 [AB007462]). They may be related to the rodent gene *Drg11*, as suggested by the BLAST searches, however, they have signatures similar to those of *Otx* and *Ptx* genes (Galliot et al. 1999), and based on the parsimony analyses, they group within the *Otd(Otx)*-containing clade. *Drg11* is a rat gene expressed in most sensory neurons and in a portion of the spinal cord (Saito et al. 1995). The potential presence of an orthologue in organisms that lack nervous organization is a provocative finding and one that requires further investigation.

HOX-Type Genes. Seven of our gene fragments are members of the HOX gene group. Such genes are best known for their roles in defining positions on the anterior–posterior axis across taxa.

Our Kelletia-5 (AY187698), Notoplana-1 (AY187699), and Neochildea-1 (AY187700) fragments are *labial* orthologues, according to both the BLAST and the parsimony analyses. Our molluscan, Kelletia-5, and flatworm, Notoplana-1, fragments are very similar to the published abalone *labial* orthologue, *Hox1_Ha* (AF327746). The acoel flatworm fragment, Neochildea-1 (AY187700), is identical in amino acid translation (two nucleic acid differences)

to the published *labial* orthologue from a parasitic trematode flatworm, *FhHbx1* (X66823).

Our molluscan fragment, Bulla-1 (AY187701), is identical in amino acid translation to both the *Drosophila proboscipedia* and a polychaete orthologue (*Pb_Nv* [AF151664]) but is only 84 and 88% similar at the nucleic acid code, respectively. To the best of our knowledge, the Bulla-1 sequence is the first report of a *pb* orthologue from the phylum Mollusca.

The fragments Lumbricus-2 (AY187702) and Urechis-2, (AY187703) are *Hox3* orthologues and we believe that the latter is the first report of any HOX gene from the phylum Echiura. According to the BLAST analysis, our scyphozoan fragment, Aurelia-1 (AY187704), is related to *AbdominalA* and is very similar to the published jellyfish sequence *scox-3* (AF124593). These two genes cluster as sister sequences in our parsimony analysis.

Pancreas/Duodenum Homeobox-*Type Genes*. Three of our molluscan gene fragments, Kelletia-1 (AY187705), Kelletia-6 (AY187706), and Kelletia-9 (AY187707), affiliate, in both the BLAST and the parsimony analyses, with the *pancreas/duodenum* (*Pdx*) genes that are known from vertebrates, a hemichordate, a cephalochordate, and an annelid. To our knowledge, these sequences represent the first report of such a gene from the phylum Mollusca. The expression of the frog and zebrafish *Pdx*-type genes suggests a conserved role in the development of the pancreas and duodenum throughout the vertebrates (Wright et al. 1989; Milewski et al. 1998). To date, there are no expression data for the other taxa from which the sequence has been isolated.

Empty Spiracles-*Type Genes*. Two of our gene fragments, Urechis-3 (AY187708) and Nutricola-1C (AY187709), are most similar to the *empty spiracles* (*ems/Emx*)-type genes and represent the first report of such genes from the superphyletic clade Lophotrochozoa. *ems/Emx*-type genes have been reported for vertebrates, cephalochordates, *Drosophila*, cnidarians, and sponges. In vertebrates, the *Emx* genes are involved with the patterning and development of portions of the central nervous system (Bonicelli et al. 1993; Pellegrini et al. 1996; Qui et al. 1996; Yoshida et al. 1997). Similarly, the *Drosophila* orthologue, *ems*, is involved in the patterning and development of the anterior head, the two posterior neuromeres of the brain, and tracheal pits/linings (Dalton et al. 1989; Hirth et al. 1995; Jones and McGinnis 1993a; Waldorf and Gehring 1992). The similar functions of this gene in these disparate taxa are provocative and suggest a conserved role in sensory organ and "head" development. This hypothesis is supported in cnidarians, where experimental conversion of posterior planula into anterior/head up-regulates the expression of the gene (Mokady et al. 2000).

Gastrulation Brain Homeobox-*Type Genes*. Three of our gene fragments, Loligo-1 (AY187710), Urechis-1C (AY187711), and Lumbricus-1 (AY187712), are closely related, based on BLAST and parsimony analyses, to the *gastrulation brain homeobox* (*Gbx*) genes that have previously been isolated from vertebrates, echinoderms, molluscs, and arthropods. Although the Loligo-1 gene is similar to the previously published squid gene *Es-Gbx2* (AF127342), the echiuran and annelidan fragments are the first report of *Gbx* orthologues from these two phyla. In vertebrates, these genes are involved in the development of the CNS, the pharyngeal endoderm, the otic placode and vesicle, and the pharyngeal arches and play a role in establishing the midbrian–hindbrain boundary (Bulfone et al. 1993; Shamim and Mason 1998; Von Bubnoff et al. 1996). In *Drosophila*, the similar *unplugged* gene is required for the formation of tracheal branches that penetrate the CNS (Chiang et al. 1995). There are no available data regarding the role of these genes in any other taxa.

Hematopoietically Expressed Homeobox-*Type Genes*. Two of our ctenophoran gene fragments, Beroe-1 (AY187713) and Beroe-2 (AY187714) BLAST as *hematopoietically expressed homeobox* genes (*Hex*). While a *Hex* orthologue has never been reported for the Ctenophora, one has been recovered from the putatively more primitive Cnidaria (*cnHex*), a finding that suggests a deeper ancestry for this gene group. In vertebrates, *Hex* genes appear to play roles in several developing organs including the thyroid, lung, liver, gallbladder, and pancreas (Bogue et al. 2000; Keng et al. 2000; Martinez et al. 2000), which are all derived from foregut endoderm. Because of their association with a portion of a germ layer, the role of *Hex* orthologues in both Cnidaria and Ctenophora would be of great interest to comparative investigators. The parsimony analyses, however, suggest that these ctenophore genes may affiliate with the *Drosophila BarH1/2* and *bsh* genes. *BarH1* and *BarH2* are involved with the patterning of distal leg segments in the fly (Kojima et al. 2000) as well as the normal development of the eye and external sensory organs (see, e. g., Hayashi et al. 1998). *Bar* genes also exist in vertebrates (see, e. g., Patterson et al. 2000) and appear to be likewise involved in aspects of visual system and central and peripheral nervous system development (see, e. g., Bulfone et al. 2000). A small subset of cells (30) in the fly brain expresses the *brain specific homeobox* (*bsh*) gene (Jones and McGinnis 1993b). To our knowledge, no other Bar-like genes have been reported (outside the vertebrates and

Drosophila). These two novel ctenophore genes change affiliation with minor modification of the parsimony analyses (data not shown), and as such, it is not clear which orthologue (*bsh* vs *Hex*) has been recovered here.

The NK-1/s59/Slouch-*Type Gene*. Our flatworm gene fragment, Himasthla-2 (AY187715), is an *NK-1*/*s59*/*slouch* orthologue and is similar to several other flatworm and nemertean *NK-1* genes (*EgHbx1* [X66817], *PnNK1* [U76499], and *LsNk* [Kmita-Cunisse et al. 1998]). In *Drosophila*, the *NK-1* gene is expressed in a discrete set of mesodermal cells (s59 derivatives) that give rise to somitic muscle tissue, a subset of neuronal cells of the CNS, and a small region of the midgut (Dohrmann et al. 1990; Knirr et al. 1999). There are no expression data from other invertebrates. The *NK-1* genes are similar to the vertebrate *SAX* genes, which are also involved in CNS development. In mouse, *Sax1* expression suggests a role in both hindbrain and spinal cord patterning, and later in development, this gene is expressed in areas of the spinal cord, hindbrain, midbrain, and forebrain (Schubert et al. 1995). These data indicate a conserved role in nervous system development.

Ladybird-*Type Gene*. Our molluscan gene fragment, Loligo-2, is related to the *ladybird* (*lbl*/*lbe*/*Lbx*) genes. While this gene group has previously been isolated from vertebrates and arthropods, Loligo-2 (AY187716) represents the first report of this gene from the phylum Mollusca. The vertebrate *Lbx* genes are involved with muscle patterning and differentiation, and they function in the development of many organs including brain, kidneys, and adrenal glands (Chen et al. 1999; Neyt et al. 2000; Uchiyama and Hanaoka 2000). The expression of the *Drosophila ladybird* genes in some mesodermal cells that correspond to heart precursors suggests an early role in heart patterning and development (Jagla et al. 1997 a, b, l998, 1999).

The ParaHOX Intermediate Neuroblasts Defective-*Type Gene*. Our molluscan gene fragment Kelletia-3 (AY187717) is related to *intermediate neuroblasts defective* (*ind*/*Gsh*), a member of the ParaHOX gene group. Orthologues of this ParaHOX gene have been identified from several basal metazoans, Placozoa and Cnidaria, and the lophotrochozoan phylum Sipunculida. In *Drosophila*, *ind* is expressed in the intermediate column cells, flanked by cells expressing *vnd* and *msh* (Weiss et al. 1998). Similarly, the vertebrate *gsh* genes (together with the *vnd* and *msh* orthologues) are expressed in corresponding ventral, intermediate, and dorsal domains during neurogenesis (reviewed by Weiss et al. 1998). These data suggest a function that is conserved across vertebrates and arthropods, but as yet, the expression of *ind* orthologues in placozoans, cnidarians, and cephalochordates remains unexplored. Such studies should reveal important information regarding the evolution of nervous system development.

Conclusion. We have exploited the high degree of sequence conservation within homeobox-containing genes to design primers for a PCR-based gene fragment isolation protocol. In a preliminary survey, we have isolated 28 gene fragments from 15 taxa that represent eight phyla. These data are the product of sequencing 18 clones per species, thus our screen does not represent an exhaustive survey, and a more thorough treatment would likely reveal numerous additional gene fragments in each species. The BLAST and parsimony analyses reveal that our gene fragments are most similar to members of the PAIRED, HOX, and ParaHox gene groups, and the *Pdx*, *ems*/*emx*, *gbx*, *Hex*, *BarH1*/*2*/*bsh*, *NK-1*/slouch, and *lbd*/*lbx* genes. These gene fragments provide the underpinnings for more detailed investigations of gene function in disparate taxa, which is a central feature of comparative evolutionary investigations. As such, this simple technique has broad utility in the field of developmental evolution.

Acknowledgments. This research was supported by National Aeronautics and Space Administration (NASA) Graduate Student Research Program Grant NGT5-550310 to S.E.L., NASA Exobiology Program Grant NAG5-7207 to D.K.J, and the Astrobiology program at UCLA. We thank V. Hartenstein for discussions and materials. We are grateful to S. Galvan, T. Huspeni, and R. Nallur for flatworm DNA extractions, to W. Hamner and L. Zeidberg for providing squid egg fingers, to W. Borgeson and the Bodega Bay Marine Lab for sending oyster specimens, to T. Huspeni for the transport of ctenophores collected by S. Anderson, and to M. Schaadt and the Cabrillo Marine Aquarium culturing staff for providing adult moon jellies. Additional technical assistance was provided by S. Galvan and V. Palchevskiy. Early versions of the manuscript were reviewed by anonymous reviewers, and we thank everyone for their insightful comments and suggestions.

References

Abzhanov A, Kaufman TC (2000) Homologs of Drosophila appendage genes in the patterning of arthropod limbs. Dev Biol 227:673–689

Agosti D, Jacobs D, DeSalle R (1996) On combining protein sequences and nucleic acid sequences in phylogenetic analysis: The homeobox protein case. Cladistics 12:65–82

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic Local Alignment Search Tool. J Mol Biol 215:403–410

Bartels JL, Murtha MT, Ruddle FH (1993) Multiple Hox/HOM-class homeoboxes in platyhelminthes. Mol Phylogenet Evol 2:143–151

Bastianello A, Ronco M, Burato PA, Minelli A (2002) Hox gene sequences from the geophilomorph centipede Pachymerium

ferrugineum (CL Koch, 1835) (Chilopoda: Geophilomorpha: Geophilidae): Implications for the evolution of the Hox class genes of arthropods. Mol Phylogenet Evol 22:155–161

Berney C, Pawlowski J, Zanietti, L (2000) Elongation factor 1-alpha sequences do not support an early divergence of the acoela. Mol Biol Evol 17:1032–1039

Bogue CW, Ganea GR, Sturm E, Ianucci R, Jacobs HC (2000) Hex expression suggests a role in the development and function of organs derived from foregut endoderm. Dev Dyn 219:84–89

Bonicelli E, Gulisano M, Broccoli V (1993) Emx and Otx homeobox genes in the developing mouse brain. J Neurobiol 24:1356–1366

Bulfone A, Puelles L, Porteus MH, Frohman MA, Martin GR, Rubenstein JLR (1993) Spatially restricted expression of Dlx-1, Dlx-2 (Tes-1), Gbx-2, and Wnt-3 in the embryonic day 12.5 mouse forebrain defines potential transverse and longitudinal segmental boundaries. J Neurosci 13:3156–3172

Bulfone A, Menguzzato E, Vani B, Marchitiello A, Gattuso C, Mariani M, Consalez GG, Martinez S, Ballabio A, Banfi S (2000) Barh 11, a gene belonging to a new subfamily of mammalian homeobox genes, is expressed in migrating neurons of the CNS. Hum Mol Genet 9:1443–1452

Carranza S, Baguña J, Riutort M (1997) Are platyhelminthes a monophyletic primitive group? An assessment using 18S rDNA sequences. Mol Biol Evol 14:485–497

Cartwright P, Dick M, Buss LW (1993) HOM/Hox type homeoboxes in the chelicerate Limulus polyphemus. Mol Phylogenet Evol 2:185–192

Chen F, Liu KC, Epstein JA (1999) *Lbx2*, a novel murine homeobox gene related to the *Drosophila ladybird* genes, is expressed in the developing urogenital system, eye, and brian. Mech Dev 84:181–184

Chiang C, Young KE, Beachy PA (1995) Control of Drosophila tracheal branching by the novel homeodomain gene unplugged, a regulatory target for genes of the bithorax complex. Development 121:3901–3912

Dalton D, Chadwick R, McGinnis W (1989) Expression and embryonic function of *empty spiracles*: A *Drosophila* homeobox gene with two patterning functions on the anterior-posterior axis of the embryo. Genes Dev 3:1940–1956

DeRosa R, Grenier JK, Andreeva T, Cook CE, Adoutte A, Akam M, Carroll SB, Balavoine G (1998) Hox genes in brachiopods and priapulids and protostome evolution. Nature 399:772–776

Dick MH, Buss LW (1994) A PCR-based survey of homeobox genes in *Ctenodrilus serratus* (Annelida:Polychaeta). Mol Phylogenet Evol 3:146–158

Dohrmann C, Azpiazo N, Frasch M (1990) A new *Drosophila* homeobox gene is expressed in mesodermal precursor cells of distinct muscles during embryogenesis. Genes Dev 4:2098–2111

Farris JS, Albert VA, Kallersjo M, Lipscomb D, Kluge AG (1996) Parsimony jackknifing outperforms neighbor-joining. Cladistics 12:99–124

Finnerty JR, Master VA, Irvine S, Kourakis MJ, Warriner S, Martindale MQ (1996) Homeobox genes in the Ctenophora: Identification of *paired*-type and Hox homologues in the atentaculate ctenophore, *Beroë ovata*. Mol Mar Bio Biotech 5:249–258

Galliot B, de Vargas C, Miller D (1999) Evolution of homeobox genes: Q50 Paired-like genes founded the Paired class. Dev Genes Evol 209:186–197

Hayashi T, Kojima T, Saigo K (1998) Specification of primary pigment cell and outer photoreceptor fates by BarH1 homeobox gene in the developing Drosophila eye. Dev Biol 200:131–145

Hirth F, Therianos S, Loop T, Gehring W, Reichert H, Furukubo-Tokunaga, K (1995) Developmental defects in brain segmentation caused by mutations of the homeobox genes *orthodenticle* and *empty spiracles* in *Drosophila*. Neuron 15:769–778

Jacobs DK, Lee SE, Dawson MN, Staton JL, Raskoff KA (1998) The history of development through the evolution of molecules: gene tress, hearts, eyes and dorsoventral inversion. In: R DeSalle, B Schierwater (eds) Molecular approaches to ecology and evolution. Birkhauser Verlag, Basel

Jagla K, Frasch M, Jagla T, Dretzen G, Bellard F, Bellard M (1997a) Ladybird, a new component of the cardiogenic pathway in Drosophila required for diversification of the heart precursors. Development 124:3471–3479

Jagla K, Jagla T, Heitzler P, Dretzen G, Bellard F, Bellard M (1997b) *ladybird*, a tandem of homeobox genes that maintain late *wingless* expression in terminal and dorsal epidermis of the *Drosophila* embryo. Development 124:91–100

Jagla T, Bellard F, Lutz Y, Dretzen G, Bellard M, Jagla K (1998) Ladybird determines cell fate decisions during diversification of Drosophila somatic muscles. Development 125:3699–3708

Jagla T, Bellard F, Vonesch J-L, Bellard M, Dastugue B, Jagla K (1999) Plasticity within the lateral somatic mesoderm of Drosophila embryos. Int Dev Biol 43:571–573

Jones B, McGinnis W (1993a) The regulation of *empty spiracles* by *Abdominal-B* mediates an abdominal segment identity function. Genes Dev 7:229–240

Jones B, McGinnis W (1993b) A new Drosophila homeobox gene, bsh, is expressed in a subset of brain cells during embryogenesis. Development 117:793–806

Katayama T, Yamamoto M, Wada H, Satoh N (1993) Phylogenetic position of acoel turbellarians inferred from partial 18S rDNA sequences. Zool Sci 10:529–536

Keng VW, Yagi H, Ikawa M, Nagano T, Myint Z, Yamada K, Tanaka T, Sato A, Muramatsu I, Okabe M, Sato M, Noguchi T (2000) Homeobox gene Hex is essential for onset of mouse embryonic liver development and differentiation of the monocyte lineage. Biochem Biophys Res Commun 276:1155–1161

Kim CB, Bailey W, Amemiya C, Miller W, Shimizu N, Ruddle FH (2000) Hox gene organization of horn shark, Heterodontus francisci. Am Zool 39:78A

Kmita-Cunisse M, Loosli F, Bièrne J, Gehring WJ (1998) Homeobox genes in the ribbonworm *Lineus sanguineus*: Evolutionary implications. Proc Natl Acad Sci USA 95:3030–3035

Knirr S, Azpiazu N, Frasch M (1999) The role of the NK-homeobox gene slouch (S59) in somatic muscle patterning. Development 126:4525–4535

Kojima T, Sato M, Kaoru S (2000) Formation and specification of distal leg elements in Drosophila by dual Bar homeobox genes, BarH1 and BarH2. Development 127:769–778

Littlewood DTJ, Rohde K, Clough KA (1999) The interrelationships of all major groups of Platyhelminthes: Phylogenetic evidence from morphology and molecules. Biol J Linn Soc 66:75–114

Martinez B, Pedro J, Clements M, Thomas P, Rodriguez T, Meloy D, Kioussis D, Beddington RSP (2000) The homeobox gene Hex is required in definitive endodermal tissues for normal forebrain, liver, and thyroid formation. Development 127:2433–2445

Milewski WM, Duguay SJ, Chan SJ, Steiner DF (1998) Conservation of PDX-1 structure, function, and expression in zebrafish. Endocrinology 139:1440–1449

Mokady O, Dick MH, Lackschewitz D, Schierwater B, Buss LW (2000) Over one-half billion years of head conservation? Expression of an *ems* class gene in *Hydractinia symbiolongicarpus* (Cnidaria:Hydrozoa). Proc Natl Acad Sci USA 95:3673–3678

Mouchel-Viehl E, Rigolot C, Gibert J-M, Deutsch JS (1998) Molecules and the body plan: The Hox genes of cirripedes (Crustacea). Mol Phylogenet Evol 9:382–389

Murtha MT, Leckman JF, Ruddle FH (1991) Detection of homeobox genes in development and evolution. Proc Natl Acad Sci USA 88:10711–10715

Neyt C, Jagla K, Thisse C, Thisse B, Haines L, Currie PD (2000) Evolutionary origins of vertebrate appendicular muscle. Nature 408:82–86

Oliver G, Vispo M, Mailhos A, Martinez C, Sosa-Pineda B, Fielitz W, Ehrlich R (1992) Homeoboxes in flatworms. Gene 121:337–342

Patterson KD, Cleaver O, Gerber WV, White FG, Krieg PA (2000) Distinct expression patterns for two Xenopus Bar homeobox genes. Dev Gene Evol 210:140–144

Pattyn A, Morin X, Cremer H, Goridis C, Brunet J-F (1997) Expression and interactions of the two closely related homeobox genes *Phox2a* and *Phox2b* during neurogenesis. Development 124:4065–4075

Pellegrini M, Mansouri A, Simeone A, Boncinelli E, Gruss P (1996) Dentate gyrus formation requires *Emx2*. Development 122: 3893–3898

Pendleton JW, Nagai BK, Murtha MT, Ruddle FH (1993) Expansion of the Hox gene family and the evolution of chordates. Proc Natl Acad Sci USA 90:6300–6304

Qui M, Anderson S, Chen S, Menese JJ, Hevner R, Kuwana E, Pedersen RA, Rubenstein JLR (1996) Mutation of the *Emx-1* homeobox gene disrupts the corpus callosum. Dev Biol 178:174–178

Ruiz-Trillo I, Riutort M, Littlewood TJ, Herniou EA, Baguña J (1999) Acoel flatworms: earliest extant bilaterian metazoans, not members of the Platyhelminthes. Science 283:1919–1923

Ruiz-Trillo I, Paps J, Loukota M, Ribera C, Jondelius U, Baguna J, Riutort M (2002) A phylogenetic analysis of myosin heavy chain type II sequences corroborates that Acoela and Nemertodermatida are basal bilaterians. Proc Natl Acad Sci USA 9:11246–11251

Saito T, Greenwood A, Qi S, Anderson DJ (1995) Identification by differential RT-PCR of a novel Paired homeodomain protein specifically expressed in sensory neurons and a subset of their CNS targets. Mol Cell NeuroSci 6:280–292

Schierwater B, Murtha M, Dick M, Ruddle FH, Buss LW (1991) Homeoboxes in cnidarians. J Exp Zool 260:413–416

Schubert FR, Fainsod A, Gruenbaum Y, Gruss P (1995) Expression of the novel murine homeobox gene Sax-1 in the developing nervous system. Mech Dev 51:99–114

Schubert M, Holland LZ, Holland ND, Jacobs DK (2000) A phylogenetic tree of the Wnt genes based on all available full-length sequences, including five from the cephalochordate amphioxus. Mol Biol Evol 17:1896–1903

Semina EV, Mintz-Hittner HA, Murray JC (2000) Isolation and characterization of a novel human paired-like homeodomain-containing transcription factor gene, *VSX1*, expressed in ocular tissues. Genomics 63:289–293

Shamim H, Mason I (1998) Expression of Gbx-2 during early development of the chick embryo. Mech Dev 76:157–159

Snow P, Buss LW (1994) Hox-type homeoboxes from *Stylaria lacustris* (Annelida:Oligochaeta). Mol Phylogenet Evol 3:360–364

Uchiyama A, Hanaoka K (2000) Expression of Lbx1 involved in the hypaxial musculature formation of the mouse embryo. J Exp Zool 286:270–279

Von Bubnoff, A, Schmidt JE, Kimelman D (1996) The Xenopus laevis homeobox gene Xgbx-2 is an early marker of anteroposterior patterning in the ectoderm. Mech Dev 54:149–160

Waldorf U, Gehring, WJ (1992) *Empty spiracles*, a gap gene containing a homeobox involved in *Drosophila* head development. EMBO J 11:2247–2259

Weiss JB, Von Ohlen T, Mellerick DM, Dressler G, Doe CQ, Scott MP (1998) Dorsoventral patterning in the Drosophila central nervous system: The intermediate neuroblasts defective homeobox gene specifies intermediate column identity. Genes Dev 12:3591–3602

Wright CVE, Schnegelsberg P, De Robertis EM (1989) *XLHbox8*: A novel Xenopus homeoprotein restricted to a narrow band of endoderm. Development 105:787–794

Xiong W-C, Okano H, Patel NH, Blendy JA, Montell C (1994) *Repo* encodes a glial-specific homeodomain protein required in the Drosophila nervous system. Genes Dev 8:981–994

Yoshida M, Suda Y, Matsuo I, Miyamoto N, Takeda N, Kuratani S, Aizawa S (1997) *Emx1* and *Emx2* functions in development of dorsal telecephalon. Development 124:101–111