Volume 16 Number 24 1988 Nucleic Acids Research

Stage- and tissue-specific expression of two homeo box genes in sea urchin embryos and adults

Gregory J.Dolecki*, Gordon Wang and Tom Humphreys

Pacific Biomedical Research Center, University of Hawaii, 41 Ahui Street, Honolulu, HI 96813, USA

Received September 7, 1988; Revised and Accepted November 11, 1988 Accession nos X13146, X13147

ABSTRACT

We report the isolation of two different homeo box genes, HB3 and HB4, from the Hawaiian sea urchin Tripneustes gratilla. DNA sequencing revealed a definitive Antennapedia (Antp) class homeo box in each gene. Southern transfer hybridizations showed the genes to be single-copy. A 5.7-kb transcript of the HB3 gene was found in ovary, testis, small intestine and gastrula poly(A)$^+$ RNA. The HB4 gene produces three transcripts. A 3.7-kb and a 4.4-kb transcript are expressed during embryogenesis. A 3.5-kb transcript appears in each of the adult tissues studied. The HB4 gene appears to be the sea urchin cognate of the Drosophila infrabdominal-7 (iab-7) gene, the mouse Hox 1.7 and Hox 3.2 genes and the Xenopus XlHbox 6 gene. An examination of Antp class homeo box genes in deuterostomes indicates that a chromosomal duplication has taken place in the evolutionary line leading to the vertebrates after the divergence of the echinoderms. Thus, the sea urchin represents the primitive condition.

INTRODUCTION

The homeo box is a 180-bp sequence which was first recognized in Drosophila development-controlling genes (1,2). It encodes a 60-aa homeo domain thought to bind DNA and regulate gene expression (3). Drosophila homeo box probes have provided a means to isolate putative development-controlling genes from other systems. Homeo box genes have been isolated by molecular cloning from the genomic DNA of humans (4,5), mice (6,7,8), rats (9), frogs (10, 11,12) and sea urchins (13,14) as well as of Drosophila.

The enterocoelous sea urchin larva probably most closely resembles the organism in which homeo box genes arose and diversified (14,15). Comparison of the function of Antp and engrailed (en) class genes in sea urchin embryos could thus provide significant insights into the original function and evolution of homeo box genes. We have previously described a sea urchin Antp

class homeo box gene expressed during embryogenesis, thus demonstrating that homeo box genes are not solely involved in the generation of a segmented body pattern (13). We have also described the only detectable sea urchin en class homeo box gene (14), concluding that the single sea urchin en class gene represents the primitive condition and that two independent duplications of en class genes have occurred, one in the protostomes and one in the deuterostomes, during evolution. In this paper, we describe two additional sea urchin Antp class homeo box genes and their transcription during embryogenesis and in adult tissues. One of these appears to be the sea urchin cognate of the Drosophila iab-7 gene.

MATERIALS AND METHODS

Sea Urchin Genomic Library Construction and Screening

Construction and screening with the Drosophila Scr probe of the recombinant phage library was described previously (13).

Radioactive Labeling of Nucleic Acid Probes

DNA probes were labelled by oligolabelling (Oligolabelling Kit, Pharmacia) according to the manufacturer's instructions using [α-^{32}P]dCTP (New England Nuclear, 3000 Ci/mmole); specific activities were approximately 1 x 10^9 dpm/μg.

Subcloning

The 3.2-kb EcoRI-SalI fragment from the λHB3 8.9-kb EcoRI insert and the 1.1-kb EcoRI-HindIII fragment from the λHB4 1.7-kb EcoRI insert were subcloned into M13mp10 and mp11 as previously described (14). The 755-bp HindIII-PvuII fragment and the 336-bp HindIII-SnaBI fragment from a deletion subclone in the HB3 series were subcloned into BluescribeM13+ (Stratagene) as were the 920-bp HindIII-EcoRI fragment and the 412-bp ClaI-PstI fragment from a deletion subclone in the HB4 series. E. coli stain JM101 was the host used for transformation (16). Recombinants carrying the desired inserts were identified by excising the inserts from small-scale plasmid or RF (replicative form) DNA preparations with the appropriate restriction endonucleases and electrophoresing the digests on agarose gels along with molecular weight standards (17). In some cases, it was necessary to subject the excised insert to a second round of digestion and electrophoresis

for identification purposes. Dale deletion subclones for nucleotide sequencing were generated from M13 recombinants (18).

Nucleotide Sequencing

M13 recombinants were sequenced by the dideoxynucleotide chain-termination method (19) using [^{35}S] deoxyadenosine 5'-(α-thio)-triphosphate as the labelled nucleotide. Both uniform and buffer gradient gels were employed (20). Nucleotide sequence analyses were performed using the GEL, GENED, SEQ, IFIND and PEP computer programs made available through the BIONET resource.

Nucleic Acid Preparation and Analysis

Isolation of the sea urchin genomic DNA used in this study has been previously described (13). Recombinant phage λHB3 and λHB4 DNAs were isolated by standard procedures (21). Rapid, small-scale isolation of recombinant plasmid DNAs and M13 RF (replicative form) DNAs were done using the alkaline lysis method (17). Single-stranded recombinant M13 DNAs to be used as templates in the dideoxy sequencing reactions were prepared as previously described (13).

Techniques for growing embryos of the Hawaiian sea urchin Tripneustes gratilla and extracting their RNA have been previously described (22). Adult tissues were collected and RNA was extracted from them as previously described (14).

Nucleic acid gel electrophoresis, transfers and hybridizations were performed as previously described (13,14,22). High stringency post-hybridization washes for Southern transfers were performed at 68°C for 20 min in 1 X SET, 0.1% SDS, 0.1% sodium pyrophosphate, pH 7.5. High stringency post-hybridizaton washes for Northern transfers were performed at 60°C for 20 min in 0.5 X SET, 0.1% SDS, 0.1% sodium pyrophosphate, pH 7.5.

RESULTS

Cloning and Sequencing of HB3 and HB4, Two New Antennapedia Class Homeo Box Genes in the Sea Urchin Genome

Eighteen bacteriophage clones were isolated after screening a Tripneustes gratilla genomic library constructed from an EcoRI partial digest of a single individual's sperm DNA with the Drosophila Sex combs reduced (Scr) homeo box probe (13). Southern transfer analysis of restriction digests revealed that three were

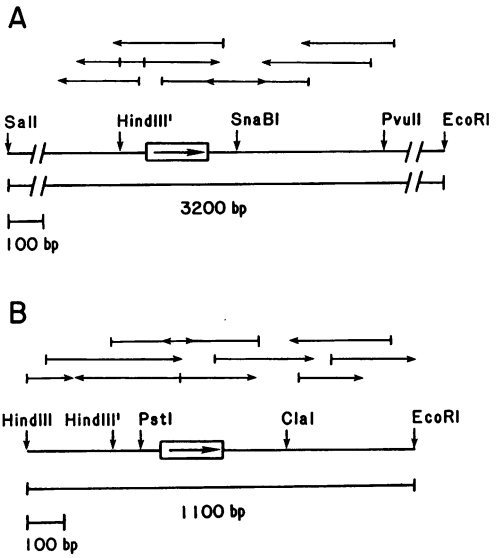

Figure 1. Restriction maps of the regions of λHB3 and λHB4 containing the sea urchin homeo boxes. Horizontal arrows on top show the sequencing strategy. All sequencing was done by the dideoxy chain termination method (19) after subcloning the restriction fragments into M13. Dale deletion subclones for nucleotide sequencing were generated from M13 recombinants (18). The rectangles indicate the homeo boxes and the arrows inside, the direction of transcription (5' to 3'). HindIII sites marked with an apostrophe (') are the ones produced by deletion subcloning which were used, where indicated, to excise DNA fragments to make probes. (A) λHB3. (B) λHB4.

The complete nucleotide sequences and the conceptual translations of the regions indicated by the sequencing strategies appear in the EMBL/GenBank Sequence Databases under the following accession numbers: HB3, X13146; HB4, X13147.

identical, each containing a 5.5- and an 8.9-kb EcoRI fragment; the 8.9-kb fragments reacted with the probe (data not shown). One of these clones, called λHB3, was selected for this study. Four other clones shared a 1.7-kb probe-reactive EcoRI fragment. Three of these were identical, each containing a 7.2-, 1.4-, 1.25-, 0.84- and 0.58-kb EcoRI fragment in addition to the 1.7-kb fragment. The fourth clone was missing the 1.4- and 1.25-kb fragments (data not shown). This clone, called λHB4, was also selected for the current study. The smallest probe-reactive fragments from λHB3 and λHB4, a 3.2-kb EcoRI-SalI fragment and a 1.1-kb EcoRI-HindIII fragment, respectively, were subcloned into

Table 1. Sequence comparisons of HB3 and HB4 with other homeo sequences

Homeo sequence (reference)	HB3 Identity (%)		HB4 Identity (%)	
	nt	aa	nt	aa
D. melanogaster				
iab-7 (24)	61	55	65	73
Antp (2)	73	80	66	68
en (25)	53	47	52	42
T. gratilla				
HB1 (13)	70	77	63	67
SU-HB-en (14)	51	47	53	42
Mouse				
Hox 1.7 (26)	59	60	76	90
Hox 1.4 (27)	69	75	64	67
Hox 3.2 (28)	63	60	80	88
En-1 (8)	54	48	51	45
En-2 (8)	52	50	51	47
X. laevis				
XlHbox 2 (12)	74	82	68	70
Xhox-1A (29)	70	77	63	63
XlHbox 6 (30,31,32)	59	60	76	87

M13 and sequenced using the strategies shown in Figure 1.

Definitive 180-bp homeo boxes, shown in Figure 2A with 42 bp of flanking nucleotides, were identified in the sequenced fragments. Within the 180 bp, HB3 is 65% identical to HB4. Table 1 shows sequence comparisons of the HB3 and HB4 homeo boxes with other *Drosophila*, sea urchin, mouse and frog homeo boxes. HB3's and HB4's identities as *Antp* class homeo boxes are established by the nucleotide sequence similarity they share with such homeo boxes, *Antp* and HB1, for example; *Antp* class homeo boxes share 60-80% nt positional identity among themselves. SU-HB-en, a typical *en* class homeo box (14), only shares approximately 50% nt positional identity with *Antp* class homeo boxes.

The amino acid sequences encoded by the HB3 and HB4 homeo boxes are shown in Figure 2B. Within the 60 aa, HB3 is 63% identical to HB4. Table 1 shows sequence comparisons of the HB3 and HB4 homeo domains with other *Drosophila*, sea urchin, mouse and

A

```
                                                    -21                      -1
                                                HB3  CAT GTT TCA TAT CTT GCA GAT
                                                HB4  A-C TGG CTC -CG GCG A-T TCC

    1                                                                        60
HB3  GGC AAA AGA GGG CGG CAG ACA TAC ACC CGC CAG CAA ACG CTG GAG TTG GAG AAG GAG TTC
HB4  --G -G- -AG AA- --C TGT C-T --- --- AAG TTT --G --- T-- --- C-- --- --A --- ---

   61                                                                       120
HB3  CAC TTC AGT CGC TAT GTT ACC CGA CGA CGT CGC TTC GAG ATC GCG CAG AGT CTT GGC CTG
HB4  -T- --T -A- ATG --C C-C --- --- GAC --C --G C-T --- --- --T -GT CTA --A A-- --C

  121                                                                       180
HB3  AGC GAA AGA CAG ATC AAG ATC TGG TTT CAG AAT AGG AGG ATG AAA TGG AAG AGG GAG CAT
HB4  -C- --G C-G --- G-- --A --- --- --C --- --C C-- --A --- --G AT- --- -A- C-- A--

  181                    201
HB3  GGG TCT AAC TGT ACC ATG ACC
HB4  CGG GCA CAA AAT TAT CCC TAA
```

B

```
                                                   -7                       -1
                                                HB3  His Val Ser Tyr Leu Ala Asp
                                                HB4  Asn Trp Leu Ser Ala Thr Ser

    1                                                                        20
HB3  Gly Lys Arg Gly Arg Gln Thr Tyr Thr Arg Gln Gln Thr Leu Glu Leu Glu Lys Glu Phe
HB4  --- Arg Lys Lys --- Cys Pro --- --- Lys Phe --- --- --- --- --- --- --- --- ---

   21                                                                        40
HB3  His Phe Ser Arg Tyr Val Thr Arg Arg Arg Arg Phe Glu Ile Ala Gln Ser Leu Gly Leu
HB4  Leu --- Asn Met --- Leu --- --- Asp --- --- Leu --- --- --- Arg Leu --- Ser ---

   41                                                                        60
HB3  Ser Glu Arg Gln Ile Lys Ile Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Arg Glu His
HB4  Thr --- --- --- Val --- --- --- --- --- --- --- --- --- --- Met --- Lys Gln Asn

   61                     67
HB3  Gly Ser Asn Cys Thr Met Thr
HB4  Arg Ala Gln Asn Tyr Pro Ter
```

Figure 2. Nucleotide sequence of the homeo box regions of the HB3 and HB4 genes and their predicted translation products. Nucleotide and amino acid matches in the HB4 sequences relative to the HB3 sequences are indicated by dashes. (A) The nucleotide sequences are shown from bp -21 (21 bp 5' of the beginning of the homeo boxes) to 222 (21 bp 3' of the homeo boxes). (B) Conceptual translations of the nucleotide sequences shown in panel A.

frog homeo domains. The amino acid sequence similarities shared by the HB3 and HB4 homeo domains and other Antp class homeo domains confirm their Antp classification; Antp class homeo domains share 60-100% aa positional identity among themselves. Similarly, SU-HB-en shares 68-80% aa positional identity with other en class homeo domains (14), but only approximately 50% aa positional identity with Antp class homeo domains.

Southern transfer analysis of restriction digests and nucle-

```
                -17                                                     -1

          HB4       Gly Arg Glu Ala Glu Glu Arg Gly His Ala Asn Trp Leu Ser Ala Thr Ser
          Hox 3.2                                                 Ile His --- Arg ---
          Hox 1.7   Pro Pro Ile Asp Pro Asn Asn Pro Ala --- --- --- --- His --- Arg ---
          XlHbox 6  Asp --- Thr His Gln Asn Asn Pro Ser --- --- --- --- His --- Arg ---
          iab-7

          1                                                               20
HB4       Gly Arg Lys Lys Arg Cys Pro Tyr Thr Lys Phe Gln Thr Leu Glu Leu Glu Lys Glu Phe
Hox 3.2   Thr --- --- --- --- --- --- --- --- --- Tyr --- --- --- --- --- --- --- --- ---
Hox 1.7   Thr --- --- --- --- --- --- --- --- --- His --- --- --- --- --- --- --- --- ---
XlHbox 6  Ser --- --- --- --- --- --- --- Ser --- Tyr --- --- --- --- --- --- --- --- ---
iab-7     Val --- --- --- --- Lys --- --- Ser --- --- --- --- --- --- --- --- --- --- ---

          21                                                              40
HB4       Leu Phe Asn Met Tyr Leu Thr Arg Asp Arg Arg Leu Glu Ile Ala Arg Leu Leu Ser Ala
Hox 3.2   --- --- --- --- --- --- --- --- --- --- --- Tyr --- Val --- --- Val --- Asn Leu
Hox 1.7   --- --- --- --- --- --- --- --- --- --- --- Tyr --- Val --- --- --- --- Asn Leu
XlHbox 6  --- --- --- --- --- --- --- --- --- --- --- His --- Val --- --- --- --- Asn Leu
iab-7     --- --- --- Ala --- Val Ser Lys Gln Lys --- Trp --- Leu --- --- Asn --- Gln Leu

          41                                                              60
HB4       Thr Glu Arg Gln Val Lys Ile Trp Phe Gln Asn Arg Arg Met Lys Met Lys Lys Gln Asn
Hox 3.2   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- Met ---
Hox 1.7   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- Ile ---
XlHbox 6  Ser --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- Leu ---
iab-7     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- Asn --- --- Asn Ser

          61
HB4       Arg Ala Gln Asn Tyr Pro Ter
Hox 3.2   Lys Glu Lys Thr Asp Lys Glu Gln Ser Ter
Hox 1.7   Lys Asp Arg Ala Lys Asp Glu Ter
XlHbox 6  Lys Asp
iab-7     Gln
```

Figure 3. Amino acid sequences of the homeo domain region of iab-7 and its deuterostome cognates. The amino acid sequences are shown from -17 (17 aa preceding the amino terminus of the homeo domains) to the end of the proteins, where data is available. Amino acid matches in other sequences relative to the sea urchin HB4 sequence are indicated by dashes. Ter, termination codon.

otide sequencing showed that the remaining 11 of the original 18 bacteriophage clones fell into an additional three groups. One represented the HB1 gene already examined (13). The other two represented new genes which will be reported on in another study (23).

HB4 Appears To Be the Sea Urchin Cognate of the Drosophila iab-7 Gene

The degree of similarity between the HB3 and the iab-7 homeo domains, 55% aa positional identity, seems low for two Antp class sequences but is in fact typical of comparisons involving the rather divergent iab-7 sequence. What is unusual is the high degree of similarity between the iab-7 homeo domain and the HB4, XlHbox 6, Hox 3.2 and Hox 1.7 homeo domains, 73%, 72%, 72% and

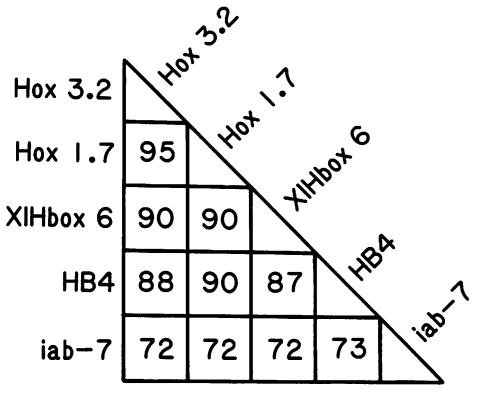

Figure 4. Similarity matrix for the amino acid sequences of the homeo domains of iab-7 and its deuterostome cognates. Numbers represent % aa positional identity, which is 100 x the number of matched aa/60 aa.

72% aa positional identity, respectively (26,28,30,31,32). Also, the HB4 homeo domain shares the highest amino acid positional identity with the Hox 1.7, Hox 3.2 and XlHbox 6 homeo domains, 90%, 88% and 87%, respectively (see Table 1).

HB4, Hox 1.7, Hox 3.2 and XlHbox 6 appear to be the deuterostome cognates of the Drosophila iab-7 gene. Figure 3 shows an alignment of the amino acid sequences of the five homeo domains and their flanking regions. Where data is available, it can be seen that the amino acid sequences preceding the homeo domains are also similar. The sequence of 8 aa immediately preceding the HB4 homeo domain is highly similar to those immediately preceding the Hox 1.7 and XlHbox 6 homeo domains. When the amino acids immediately preceding just the latter two homeo domains are compared, the sequence similarity extends to at least 12 aa with only a single variation. The amino acids following the homeo domains do not show the same level of sequence similarity. Figure 4 shows a similarity matrix for the five homeo domain sequences.

HB3 and HB4 Restriction Fragment Length Polymorphism in the Tripneustes gratilla Genome

To establish whether the 8.9- and 1.7-kb homeo box-containing EcoRI fragments from λHB3 and λHB4, respectively, derive from single-copy genes, Southern transfers of EcoRI- and HindIII-di-

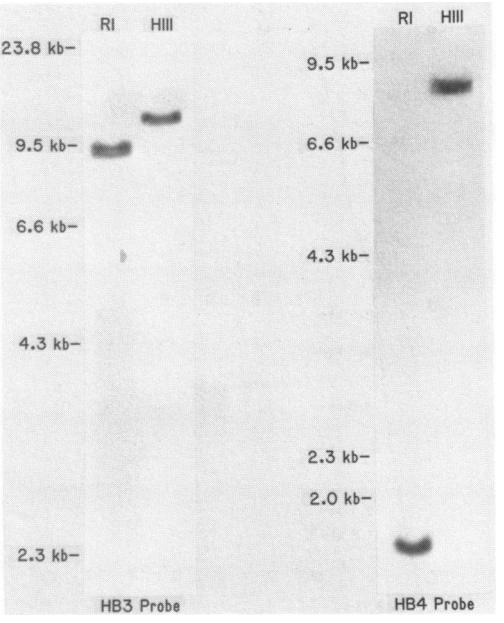

Figure 5. Hybridization of HB3 and HB4 homeo box probes to Southern transfers of T. gratilla genomic DNA. DNA from the individual used to construct the genomic library was digested to completion with EcoRI or HindIII, electrophoresed on a 0.8% agarose gel and transferred to nitrocellulose; 15 µg of DNA were loaded in each lane. Restriction fragments used to make the probes are described in MATERIALS AND METHODS. Post-hybridization washes were high stringency. Filters were exposed to Kodak XAR-5 film for 1 to 5 days. RI = EcoRI, HIII = HindIII.

gested genomic DNA from the individual used to construct the genomic library were reacted with a ^{32}P-labelled 755-bp HindIII-PvuII homeo box-containing fragment from a deletion subclone in the HB3 series and a ^{32}P-labelled 412-bp ClaI-PstI homeo box-containing fragment from a deletion subclone in the HB4 series (see Figure 1). As shown in Figure 5, after high stringency post-hybridization washes, the HB3 probe reacted with an 8.9- and a 9.2-kb EcoRI fragment and a 10.7-kb HindIII fragment; the HB4 probe reacted with a single 1.7-kb EcoRI fragment and a single 8.4-kb HindIII fragment. The bands at 8.9 and 1.7 kb in the EcoRI lanes on the autoradiograph in Figure 5 correlate with the sizes of the probe-reactive fragments cloned in λHB3 and λHB4.

Our previous study of restriction fragment length polymorphism in the T. gratilla genome detected by the Drosophila Scr

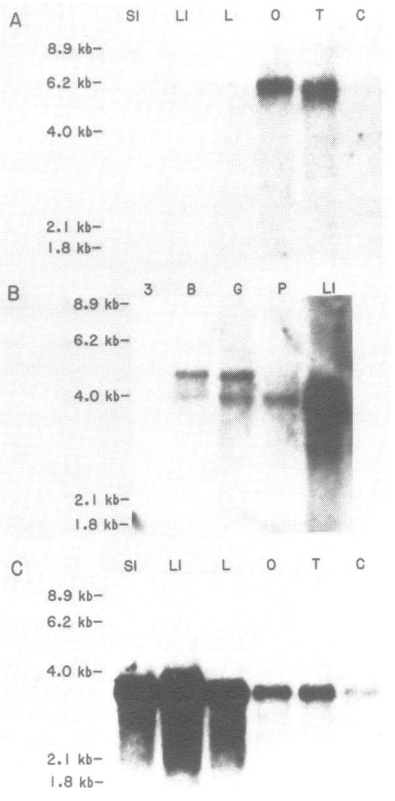

Figure 6. Northern transfers of poly(A)$^+$RNA from sea urchin embryos and adult tissues reacted to HB3 and HB4 homeo box probes. Poly(A)$^+$RNA, 2 µg per lane, from the developmental stages and adult tissues indicated was electrophoresed on denaturing 1% formaldehyde-agarose gels, transferred to nitrocellulose and hybridized to HB3 and HB4 homeo box probes. Denatured HindIII-digested λ DNA was used for size markers. Post-hybridization washes were high stringency. Filters were exposed to Kodak XAR-5 film for 2 weeks at -70° C with intensifying screens. Developmental stages: 3, 3 hr; B, blastula; G, gastrula; P, pluteus. Adult tissues: C, coelomocytes; T, testes; O, ovaries; L, Aristotle's lanterns; LI, large intestines; SI, small intestines. (A) The HB3 probe used was a 336-bp homeo box-containing HindIII-SnaBI fragment from a deletion subclone (see map in Figure 1A). This probe produces the same results with Southern transfers as the larger 755-bp HindIII-PvuII probe but eliminates high background hybridization produced by the larger probe with Northern transfers (data not shown). (B) The HB4 probe used was the 412-bp ClaI-PstI homeo box-containing fragment used for Southern transfer hybridizaton. (C) The HB4 probe used was a 920-bp homeo box-containing HindIII-EcoRI fragment from a deletion subclone (see

map in Figure 1B). The probe used in panel B produces the same reaction with adult tissue RNA as the one used in this panel (see panel B, lane LI).

Although separate transfers were made for panels A and B, the RNA on each transfer is from the same sample preparation.

probe indicated that contrary to predictions based on an average single-copy sequence diversity of 4%, the majority of the hybridizing fragments were the same in the EcoRI- and HindIII-digested DNA of five different individuals, one of which was the same individual used to construct the genomic library (13). This is significantly less polymorphism than is observed around other sea urchin genes that have been examined (22,33). The only major bands exhibiting extensive restriction fragment length polymorphism were the largest ones in the EcoRI and HindIII digests. The present study enables us to identify the HB3 homeo box-containing fragments as the ones which exhibited the most extensive restriction fragment length polymorphism in our previous study. We conclude that the bands at 8.9 and 9.2 kb in the EcoRI lane in Figure 5A represent different alleles of the HB3 gene (22,33).

The single 1.7-kb EcoRI fragment and the single 8.4-kb HindIII fragment reacting with the HB4 probe (Figure 5B) indicate that the HB4 gene is single-copy. These fragments showed only one restriction fragment length variant out of 10 alleles in our previous study. Thus, the HB4 homeo box is one of the sea urchin homeo boxes embedded in highly conserved DNA (13).

Transcription of the HB3 and HB4 Genes

To investigate the transcription of the HB3 and HB4 genes in sea urchin embryos and adult tissues, Northern transfers of 2 μg of poly(A)$^+$ RNA from 3-hr, blastula-, gastrula- and pluteus-stage embryos and from coelomocytes, testes, ovaries, Aristotle's lanterns, large intestines and small intestines were reacted with ^{32}P-labelled HB3 and HB4 homeo box probes. High stringency post-hybridization washes were performed.

As shown in Figure 6A, a 5.7-kb HB3 transcript was found in ovary and testis poly(A)$^+$ RNA. A very faint reaction occurred at 5.7 kb in the small intestine lane, but it is not visible in the photograph. An equally faint reaction occurred at 5.7 kb in the gastrula lane on Northern transfers of embryo poly(A)$^+$ RNA (data not shown).

Two HB4 transcripts were detected in embryo poly(A)$^+$ RNA as shown in Figure 6B. A 3.7-kb transcript first appears at the blastula stage and then increases in abundance to a maximum at the gastrula and pluteus stages. A 4.4-kb transcript, somewhat more abundant than the 3.7-kb transcript, also first appears at the blastula stage, increases in abundance to a maximum at the gastrula stage but then becomes almost undetectable by the pluteus stage.

Figure 6C shows a single 3.5-kb HB4 transcript in each of the adult tissue poly(A)$^+$ RNAs. It is most abundant in the large intestine RNA, about half as abundant in the small intestine and Aristotle's lantern RNA, about half again as abundant in the ovary and testis RNA and barely detectable in the coelomocyte RNA. As seen in Figure 6B, a lane of large intestine poly(A)$^+$ RNA run on the same gel as the embryo RNA shows that the adult tissue transcript is many times more abundant in the large intestine RNA than either of the embryo transcripts are in the RNA of any stage investigated. Also, although it is not obvious from the exposure used in Figure 6B, the 3.5 kb adult tissue transcript is distinctly different in size from the 3.7 kb embryo transcript.

DISCUSSION

The data presented here describe two new *Antp* class sea urchin homeo box genes, HB3 and HB4, and provide some information about their developmental regulation and adult tissue-specific expression.

The sea urchin HB1 homeo box gene like *Drosophila* homeo box genes is expressed as multiple mRNAs during specific periods of development (13). HB3 is unusual in this regard; a single, barely detectable 5.7-kb transcript appears only in gastrula-stage poly(A)$^+$ RNA. This may be the result of transcription at extremely low levels or in a very restricted cell population not adequately represented in our samples. Similarly, we were unable to detect any transcripts of the sea urchin *en* class SU-HB-en gene during embryogenesis (14).

Again like SU-HB-en which is expressed predominantly in Aristotle's lanterns, HB3 transcription is tissue-specific in

adult sea urchins. Transcripts of 5.7-kb are clearly present in the poly(A)$^+$ RNA from the ovaries and testes of animals that had been induced to shed their mature gametes by KCl injection. Except for a barely detectable 5.7-kb band in the small intestine lane, no other adult tissue showed any evidence of HB3 transcription. Perhaps HB3 transcripts are germ cell-specific. In mice, transcripts of the Hox 1.4 gene are testis-specific (27). Moreover, they appear to be limited to germ cells that have entered meiosis (34). In Xenopus, high levels of XlHbox 2 gene transcripts are found in oocytes (12). HB3 is probably not the sea urchin cognate of the Hox 1.4 gene or the XlHbox 2 gene, however; an analysis of all published homeo domains shows that the HB3 homeo domain is more similar to many other homeo domains than it is to the Hox 1.4 or XlHbox 2 homeo domains.

HB4 transcription is more typical of homeo box genes; two developmentally regulated transcripts are produced. The 3.7-kb transcript increases in abundance during sea urchin embryogenesis while the 4.4-kb transcript increases and then decreases in abundance. A third 3.5-kb transcript is the only one which appears in adults, but it does not seem to be tissue-specific. Although it is most abundant in the poly(A)$^+$ RNA from large intestines, it is present at a significant level in the poly(A)$^+$ RNA from five of the six adult tissues we investigated. Only coelomocyte poly(A)$^+$ RNA has a barely detectable level of the 3.5-kb transcript.

The high degree of sequence similarity shared by the iab-7, HB4, XlHbox 6, Hox 3.2 and Hox 1.7 homeo domains leads us to believe that HB4, Hox 1.7, Hox 3.2 and XlHbox 6 may be the deuterostome cognates of the Drosophila iab-7 gene. The similarity of the amino acid sequences immediately preceding the four deuterostome homeo domains supports our conclusion; unfortunately, the amino acid sequence preceding the iab-7 homeo domain is not available for comparison. As expected, the four deuterostome homeo domains are the most similar since they are the most closely related (see Figure 4). However, as with the Drosophila Deformed (Dfd) gene and its cognates in frogs, mice and man (29,35,36,37), isolation and sequencing of complete cDNA clones of the tran-

scripts of these genes will be needed to confirm their relationship.

An examination of the Antp class homeo box genes in the deuterostome branch of the evolutionary tree reveals not only cognates of Drosophila genes, the most well-studied protostome genes, indicating that these sequences arose in a common ancestor (13,14), but also intra-species cognates. This sort of homeo box gene duplication is not seen among the Antp class genes in Drosophila. In mice at least, where the question can be easily addressed; these intra-species cognates map to different chromosomes (38). As described above, the mouse genes Hox 1.7 and Hox 3.2 appear to be closely related; Hox 1.7 is on chromosome 6 and Hox 3.2, on chromosome 15 (26,28). Also, the two Xenopus genes XHox-36 and XlHbox 2, the two mouse genes Hox 1.1 and Hox 2.3 and the human gene HHO.c1 can be grouped together into a distinct family (12); Hox 1.1 is on mouse chromosome 6 (39) and Hox 2.3, on mouse chromosome 11 (40).

Ruddle (38) presents evidence that chromosomal duplication is responsible for the gene duplication in the deuterostomes. However, as in Drosophila, there is no evidence for homeo box gene duplication in sea urchins. We have found only a single iab-7 cognate, HB4, in Tripneustes and no two clones that we have isolated as a result of screening our genomic library for homeo box sequences have shown a higher percent positional identity in deduced amino acid sequence alignments with each other than with the "ancestral" Antp gene (13, data not shown). Based on our data, the simplest explanation is that the proposed chromosomal duplication must have taken place in the evolutionary line leading to the vertebrates after the divergence of the echinoderms.

This explanation is in good agreement with the conclusions from our study of en class genes that the sea urchin represents the primitive condition with a single en class gene and that two independent en class gene duplication events took place, one in the protostome line and one in the deuterostome line after the divergence of the echinoderms (14). It is interesting to note that the two mouse en class genes, En-1 and En-2, are on different chromosomes in keeping with the proposed chromosomal duplication event in deuterostomes (8) while the two Drosophila en class

genes, en and invected (inv), are within 17 kb of one another on the same chromosome (41), suggesting a tandem duplication event.

ACKNOWLEDGEMENTS

We appreciate the expert technical help of Dawn Haeckel and Bert Fukunaga. This work was supported by grants from the National Institutes of Health to G.J.D. (HD22483) and T.H. (HD06574). The BIONET resource is supported by a grant from the National Institutes of Health Division of Research Resources (RR01685).

*To whom correspondence should be addressed at Kewalo Marine Laboratory, 41 Ahui Street, Honolulu, HI 96813, USA

REFERENCES

1. McGinnis, W., Garber, R.L., Wirz, J., Kuroiwa, A. and Gehring, W.J. (1984) Cell 37, 403-408.
2. Scott, M.P. and Weiner, A.J. (1984) Proc. Natl. Acad. Sci. USA 81, 4115-4119.
3. Laughon, A. and Scott, M.P. (1984) Nature 318, 25-31.
4. Levine, M., Rubin, G.M. and Tjian, R. (1984) Cell 38, 667-673.
5. Simeone, A., Mavilio, F., Bottero, L., Giampaolo, A., Russo, G., Faiella, A., Boncinelli, E. and Peschle, C. (1986) Nature 320, 763-765.
6. Hart, C.P., Awgulewitsch, A., Fainsod, A., McGinnis, W. and Ruddle, F.H. (1985) Cell 43, 9-18.
7. Meijlink, F., de Laaf, R., Verrijzer, P., Destree, O., Kroezen, V., Hilkens, J. and Deschamps, A. (1987) Nucleic Acids Res. 15, 6773-6786.
8. Joyner, A.L. and Martin, G.R. (1987) Genes Dev. 1, 29-38.
9. Falzon, M., Sanderson, N. and Chung, S. (1987) Gene 54, 23-32.
10. Carrasco, A.E., McGinnis, W., Gehring, W.J. and De Robertis, E.M. (1984) Cell 37, 409-414.
11. Muller, M.M., Carrasco, A.E. and De Robertis, E.M. (1984) Cell 39, 157-162.
12. Wright, C.V.E., Cho, K.W.Y., Fritz, A., Burglin, T.R. and De Robertis, E.M. (1987) EMBO J. 6, 4083-4094.
13. Dolecki, G.J., Wannakrairoj, S., Lum, R., Wang, G., Riley, H.D., Carlos, R., Wang, A. and Humphreys, T. (1986) EMBO J. 5, 925-930.
14. Dolecki, G.J. and Humphreys, T. (1988) Gene 64, 21-31.
15. Hyman, L.H. (1951) The Invertebrates, Vol. II, pp. 18-28. McGraw-Hill, New York.
16. Messing, J. (1983) Methods Enzymol. 101, 20-79.
17. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor.
18. Dale, R.M., McClure, B.A. and Houchins, J.P. (1985) Plasmid 13, 31-41.
19. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.

20. Biggin, M.D., Gibson, T.J. and Hong, G.F. (1983) Proc. Natl. Acad. Sci. USA 80, 3963-3965.
21. Blattner, F.R., Williams, B.G., Blechl, A.E., Denniston-Thompson, K., Faber, H.E., Furlong, L., Grunwald, D.J., Kiefer, D.O., Moore, D.D., Schumm, J.W., Sheldon, E.L. and Smithies, O. (1977) Science 196, 161-169.
22. Fregien, N., Dolecki, G.J., Mandel, M. and Humphreys. T. (1983) Mol. Cell. Biol. 3, 1021-1031.
23. Wang, G., Dolecki, G.J. and Humphreys, T. in preparation.
24. Regulski, M., Harding, K., Kostriken, R., Karch, F., Levine, M. and McGinnis, W. (1985) Cell 43, 71-80.
25. Fjose, A., McGinnis, W. and Gehring, W.J. (1985) Nature 313, 284-289.
26. Rubin, M.R., King, W., Toth, L.E., Sawczuk, I.S., Levine, M.S., D'Eustachio, P. and Nguyen-Huu, M.C. (1987) Mol. Cell. Biol. 10, 3836-3841.
27. Wolegmuth, D.J., Engelmyer, E., Duggal, R.N., Gizang-Ginsberg, E., Mutter, G.L., Ponzetto, C., Viviano, C. and Zakeri, Z.F. (1986) EMBO J. 5, 1229-1235.
28. Breier, G., Dressler, G.R. and Gruss, P. (1988) EMBO J. 7, 1329-1336.
29. Harvey, R.P., Tabin, C.J. and Melton, D.A. (1986) EMBO J. 5, 1237-1244.
30. Sharpe, C.R., Fritz, A., De Robertis, E.M. and Gurdon, J.B. (1987) Cell 50, 749-758.
31. De Robertis, E.M., Burglin, T.R., Fritz, A., Wright, C.V.E., Jegalian, B., Schnegelsberg, P., Bittner, D., Morita, E., Oliver, G. and Cho, K.W.Y. (1988) In Grala, J. (ed), UCLA Symposium, Alan R. Liss, Inc., New York, Vol. 95, pp. 1-9.
32. Cho, K.W.Y., Wright, C.V.E. and De Robertis, E.M. personal communication.
33. Posakony, J.W., Flytzanis, C.N., Britten, R.J. and Davidson, E.H. (1983) J. Mol. Biol. 167, 361-389.
34. Wolgemuth, D.J., Viviano, C.M., Gizang-Ginsberg, E., Frohman, M.A., Joyner, A.L. and Martin, G.R. (1987) Proc. Natl. Acad. Sci. USA 84, 5813-5817.
35. Regulski, M., McGinnis, N., Chadwick, R. and McGinnis, W. (1987) EMBO J. 6, 767-777.
36. Laughon, A., Carroll, S.B., Storfer, F.A., Riley, P.D. and Scott, M.P. (1985) Cold Spring Harbor Symp. Quant. Biol. 50, 253-262.
37. Mavilio, F., Simeone, A., Giampaolo, A., Faiella, A., Zappavigna, V., Acampora, D., Poiana, G., Russo, G., Peschle, C. and Boncinelli, E. (1986) Nature 324, 664-667.
38. Ruddle, F.H. (1987) In Proceedings of the Society for Study of Human Biology Symposium: Physiology of Growth, Cambridge University Press, Cambridge, In Press.
39. McGinnis, W., Hart, C.P., Gehring, W.J. and Ruddle, F.H. (1984) Cell 38, 675-680.
40. Rabin, M., Hart, C.P., Ferguson-Smith, A., McGinnis, W., Levine, M. and Ruddle, F. (1985) Nature 314, 175-178.
41. Poole, S.J., Kauvar, L.M., Drees, B. and Kornberg, T. (1985) Cell 40,37-43.