***Xenopus* homeobox-containing cDNAs expressed in early development**

Andreas Fritz and Eddy M.De Robertis

Department of Biological Chemistry, University of California, Los Angeles, CA 90024-1737, USA

Received August 31, 1987; Revised and Accepted January 6, 1988

ABSTRACT

We report the isolation of six different homeobox-containing genes in *Xenopus laevis* which are expressed during early embryogenesis. cDNA clones of all of them were partially sequenced including two from previously isolated *Xenopus* genomic loci and found to contain homeodomains which share a high degree of homology with the Antennapedia protein (50 to 59 out of 60 amino acids). We find a short region of homology (consensus Ile Tyr Pro Trp Met) in four of the cDNAs, which is also present in Antennapedia and that may correspond to a bend region preceding the homeodomain. Northern blots have been performed to show the transcriptional pattern through early frog embryogenesis. Three of the genes are expressed only during a very narrow period of embryogenesis, reinforcing the view that homeobox genes are developmentally regulated.

INTRODUCTION

The homeobox is a 180 nucleotide protein coding region conserved between a number of *Drosophila* development-controlling genes (1,2). We have previously used this sequence to identify and isolate the first vertebrate homeobox-containing gene from a *Xenopus laevis* genomic library (3). Similar genes have subsequently been identified in a variety of vertebrates, including man (1,4,5,6,7,8). We now report the isolation of *Xenopus* cDNA clones corresponding to six different homeobox-containing genes which have been named X1Hbox1,2,3,4,5 and 6, respectively, X1Hbox meaning *Xenopus laevis* homeobox (X1Hbox1 corresponds to previously isolated gene AC1, 3; and X1Hbox2 to MM3, 5). We present Northern blots to determine the developmental profile of transcription for the new genes X1Hbox3, 4 and 5. The sixth gene isolated during this screen, which has the interest of being expressed in the posterior neural plate, is described elsewhere (9). The homeobox sequence of X1Hbox4 is identical to a *Xenopus* clone called Xhox1B (7). A representative cDNA of each of those classes was chosen and partially sequenced. We find that three out of the five classes (X1Hbox1, 2 and 4) share a high degree of homology outside of the homeobox domain with homeobox-containing genes reported in man and

mouse (10,11,12,13). We furthermore find a region of homology within the 10-20 amino acids preceding the homeobox (consensus Ile Tyr Pro Trp Met), which is also present in the Antennapedia (Antp) protein (14,15,16).

RESULTS

Isolation of Frog Homeobox-containing cDNAs

Because we are interested in early *Xenopus* development, we decided to isolate cDNA clones that are expressed in mid-gastrula embryos. A gastrula cDNA library (stage 12, amplified once) in the vector λ gt 10 was generously provided by Richard Harland (U. of California, Berkeley). Screening of 10^6 λ recombinants was carried out under low stringency conditions (1) using a mixed nick-translated probe. The probe consisted of a 600 bp fragment of Antp (1), a 118 bp fragment of the genomic X1Hbox1 clone (3) and a 250 bp fragment of X1Hbox2 (5), all of which contain either part of (X1Hbox1) or the entire homeobox region.

Thirty-six hybridizing clones were isolated, purified and subcloned into plasmid vectors (see Materials and Methods). The clones were classified into six groups by high stringency hybridizations using nick-translated inserts from the plasmid vectors and hybridizing them back to the original λ clones which had been dot-blotted onto nitrocellulose filters. The classification was then confirmed using synthetic oligonucleotides directed against sequences in the homeobox-region which were identical in the X1Hbox1 and X1Hbox2 genes (see Materials and Methods).

The longest member of each class was chosen to establish a restriction map of the transcript as shown in Fig. 1A. Figure 1B shows the number of isolates obtained for each of the six classes.

The *Xenopus* Homeoboxes Show a High Degree of Homology to the Antennapedia Homeobox

Figure 2A shows the nucleotide sequences of all five X1Hbox-genes starting 5' to the homeobox-region and ending 3' to the homeobox at the putative translation stop. Figure 2B shows the deduced protein sequences corresponding to the nucleotide sequences in Fig. 2A.

The frog homeobox-domains all show a high degree of homology to each other and have not diverged considerably from the Antp domain. In fact, the frog genes show a degree of homology to Antp which is comparable to the homology shared by the Antp, Ubx and ftz homeo domains between themselves. Therefore, we refer to the frog genes as "Antp-like" (see also 17). Apparently we have been unable, with the probes and conditions used, to obtain

A

B

X1Hbox	# of clones	reference names (*)
1:	11	p10, p12, p51, p15, p18, pS23, p25, p30, p5, p34, and p41
2:	9	p52, pS13, p4, p16, p29, p2, p8, p9, and p11
3:	10	pS47, p6, p7, p17, p19, p24, p35, p40, p49, and p14
4:	3	p22, p28, and p3
5:	2	p26 and p48
6:	1	p1

(*): clones ordered in decreasing length

Figure 1:

- A: Restriction map of longest representative of each X1Hbox class: X1Hbox1: p10; X1Hbox2: p52; X1Hbox3: p7; X1Hbox4: p28; X1Hbox5: p26; X1Hbox6: p1. Restriction enzymes are: A : ApaI; B : BamHI; Bg: BglII; E : EcoRI; K : KpnI; P : PstI; Sa : SacI; S : SmaI; X : XbaI. The position of the homeobox is indicated by a box.
- B: Table of cDNA clones obtained

homeobox-containing genes of groups with lower homologies, e.g.

"engrailed-like" transcripts, which have been successfully cloned from mouse (18).

The overall structure of the frog transcripts is very similar to that of the *Drosophila* genes. In all cases where we know the genomic structure, i.e. X1Hbox1,2,4 and 6 (refs. 3,5,7: and our unpublished data), the homeodomain is located in the last exon, as it is in most *Drosophila* genes and is close to the COOH-terminus of the protein. The putative translation stops are between

A	X1Hbox1	-60	CAG ATC TAC CCA TGG ATG CAG CGC ATG AAC TCC CAC AGT GGC GTG GGC TAC GGG ACG GAC	-45	---	-30	---	-15	---
	X1Hbox2	---	CTG CAC AAT GAG GCT AAC TTG AGA ATC TAC CCC TGG ATG AGG AGC GCA GGG TGG GAC	---	---	---	---	---	---
	X1Hbox3	---	---	---	TCT CCC AGG TGC CCT GGT TTT CTT TAG AGA AAA GGT CCG GAC	---	---	---	---
	X1Hbox4	---	CCC CAG ATC TTC CCC TGG ATG AGA AAA CTG CAC ATT AAC CAC GAC ATG GCA GGT CCA GAT	---	---	---	---	---	---
	X1Hbox5	---	GAA TTC CGA CAA ATA TAT CCT TGG ATG ACC AAA CTG CAC ATG AGC CAT GAA ACG GAC	---	---	---	---	---	---
	X1Hbox1	1	AGG AGG AGA GGA CGC CAG ATC TAT TCC CGT TAC CAA ACT CTG GAG CTG GAG AAA GAA TTT	16	AGG AAG AGG GGT CGC CAG ACC TAC ACA AGG TAC CAG ACC CTG GAG CTG GAG AAG GAG TTT	31	AGG AAG AGG GGT CGC CAG ACT TAC ACC CGT TAC CAG ACT CTA GAG CTG GAG AAA GAA TTT	46	GGG AAG AGG GCG AGG ACT GCC TAC ACC CGG TAC CAG ACG CTG GAG CTG GAG AAG GAG TTC
	X1Hbox2	---	AGG AAG AGG GGT CGC CAG ACC TAC ACA AGG TAC CAG ACC CTG GAG CTG GAG AAG GAG TTT	---	---	---	---	---	---
	X1Hbox3	---	AGG AAG AGG GGT CGC CAG ACT TAC ACC CGT TAC CAG ACT CTA GAG CTG GAG AAA GAA TTT	---	---	---	---	---	---
	X1Hbox4	---	GGG AAG AGG GCG AGG ACT GCC TAC ACC CGG TAC CAG ACG CTG GAG CTG GAG AAG GAG TTC	---	---	---	---	---	---
	X1Hbox5	---	GGG AAG CGT TCC CGG ACA AGT TAC ACT CGA TAC CAG ACC CTG GAG CTG GAG AAG GAA TTC	---	---	---	---	---	---
	X1Hbox1	61	CAC TTC AAT CGC TAC CTG ACC CGG CGC AGG AGG ATC GAG ATC GCC AAT GCA CTT TGT CTC	76	CAC TTC AAT CGC TAC CTG ACC CGG CGG AGA CGC ATA GAA ATC GCT CAC ACT CTG TGT CTC	91	CAC TTC AAT CGC TAC CTA ACC CGC AGG AGG AGA ATA GAA ATC GCA CAC GCT TGT CTC	106	CAT TTC AAT AGG TAC CTG ACC ACG AGG AGG AGG ATC GAG ATC GCC CAC ACC TTG TGC CTC
	X1Hbox2	---	CAC TTC AAT CGC TAC CTG ACC CGG CGG AGA CGC ATA GAA ATC GCT CAC ACT CTG TGT CTC	---	---	---	---	---	---
	X1Hbox3	---	CAC TTT AAT CGC TAC CTA ACC CGC AGG AGG AGA ATA GAA ATC GCA CAC GCT TGT CTC	---	---	---	---	---	---
	X1Hbox4	---	CAT TTC AAT AGG TAC CTG ACC ACG AGG AGG AGG ATC GAG ATC GCC CAC ACC TTG TGC CTC	---	---	---	---	---	---
	X1Hbox5	---	CAC TTT AAT AGA TAC CTG ACT CGC CGA AGG CGC ATA GAG ATC GCC AAC AAC CTC TGT CTA	---	---	---	---	---	---
	X1Hbox1	121	ACA GAG CGG CAG ATC AAA ATC TGG TTC CAG AAC AGG AGG ATG AAA TGG AAA AAG GAG AGC	136	ACC GAG AGA CAA ATC AAA ATC TGG TTC CAG AAC CGC AGG ATG AAA TGG AAA AAG GAA AAC	151	ACT GAA AGG CAG ATC AAA ATC TGG TTC CAG AAC CGG AGG ATG AAA TGG AAG AAA GAA CAC	166	TCC GAG CGG CAG ATT AAA ATT TGG TTC CAG AAC CGG AGG ATG AAG TGG AAA AAG GAC AAC
	X1Hbox2	---	ACA GAG CGG CAG ATC AAA ATC TGG TTC CAG AAC AGG AGG ATG AAA TGG AAA AAG GAG AGC	---	---	---	---	---	---
	X1Hbox3	---	ACT GAA AGG CAG ATC AAA ATC TGG TTC CAG AAC CGC AGG ATG AAA TGG AAG AAA GAA CAC	---	---	---	---	---	---
	X1Hbox4	---	TCC GAG CGG CAG ATT AAA ATT TGG TTC CAG AAC CGG AGG ATG AAG TGG AAA AAG GAC AAC	---	---	---	---	---	---
	X1Hbox5	---	AAC GAG AGG CAG ATC AAA ATT TGG TTC CAG AAC CGA CGG ATG AAA TGG AAG AAA GAC ACA	---	---	---	---	---	---
	X1Hbox1	181	AAC CTC TCA TCT ACC CTT OCT GGT GGC ACT GGG GCA GCA GCT GAC AGT TTA GCC GAG GGT	196	AAA GCG TCC AGC CCT TCC TCT AAC AGC CAA GAA AAG CCG GAG ACT GAG GAG GAG GAG	211	AAG GAG GAA AGT GAC CAG ACC CCC GAC GCT GGG GAG GAA TCC ACT GCA CCC ACT ACT ACT	226	AAG CTG AAA AGC ATG AGC CTC GCC ACG GGC AGC AGC GCC TTC CAG CCC TGA
	X1Hbox2	---	AAC CTC TCA TCT ACC CTT OCT GGT GGC ACT GGG GCA GCA GCT GAC AGT TTA GCC GAG GGT	---	---	---	---	---	---
	X1Hbox3	---	AAG GAG GAA AGT GAC CAG ACC CCC GAC GCT GGG GAG GAA TCC ACT GCA CCC ACT ACT ACT	---	---	---	---	---	---
	X1Hbox4	---	AAG CTG AAA AGC ATG AGC CTC GCC ACG GGC AGC AGC GCC TTC CAG CCC TGA	---	---	---	---	---	---
	X1Hbox5	---	AAA GTC AAG AGC AAA GAC TCC ATG TAA	---	---	---	---	---	---
	X1Hbox1	241	AAG GAG GAA AAG CAA GAG GAC TCA GAG GGA CAA GGC AAA GAG TGA	256	GAG GAG GAA TGA	271	GCA GAT GAT GAT AAA GAT AAA GAA TAA		
	X1Hbox2	---	GAG GAG GAA TGA	---	---	---	---		
	X1Hbox3	---	GCA GAT GAT GAT AAA GAT AAA GAA TAA	---	---	---	---		

B	X1Hbox1	-21	Gln Ile Tyr Pro Trp Met Gln Arg Met Asn Ser His Ser Gly Val Gly Tyr Gly Thr	-1	Asp
	X1Hbox2	---	Leu His Asn Glu Ala Asn Leu Arg Ile Tyr Pro Trp Met Arg Ser Ala Gly Ser	---	Asp
	X1Hbox3	---	---	---	Asp
	X1Hbox4	---	Pro Gln Ile Phe Pro Trp Met Arg Lys Leu His Ile Asn His Asp Met Ala Gly Pro	---	Asp
	X1Hbox5	---	--- Glu Phe Arg Gln Ile Tyr Pro Trp Met Thr Lys Leu His Met Ser His Gly Thr	---	Asp
	X1Hbox1	1	Arg Arg Arg Gly Arg Gln Ile Tyr Ser Arg Tyr Gln Thr Leu Glu Leu Glu Lys Glu Phe	20	Phe
	X1Hbox2	---	Arg Lys Arg Gly Arg Gln Thr Tyr Thr Arg Tyr Gln Thr Leu Glu Leu Glu Lys Glu Phe	---	Phe
	X1Hbox3	---	Arg Lys Arg Gly Arg Gln Thr Tyr Thr Arg Tyr Gln Thr Leu Glu Leu Glu Lys Glu Phe	---	Phe
	X1Hbox4	---	Gly Lys Arg Ala Arg Thr Ala Tyr Thr Arg Tyr Gln Thr Leu Glu Leu Glu Lys Glu Phe	---	Phe
	X1Hbox5	---	Gly Lys Arg Ser Arg Thr Ser Tyr Thr Arg Tyr Gln Thr Leu Glu Leu Glu Lys Glu Phe	---	Phe
	X1Hbox1	21	His Phe Asn Arg Tyr Leu Thr Arg Arg Arg Arg Ile Glu Ile Ala Asn Ala Leu Cys Leu	40	Leu
	X1Hbox2	---	His Phe Asn Arg Tyr Leu Thr Arg Arg Arg Arg Ile Glu Ile Ala His Thr Leu Cys Leu	---	Leu
	X1Hbox3	---	His Phe Asn Arg Tyr Leu Thr Arg Arg Arg Arg Ile Glu Ile Ala His Ala Leu Cys Leu	---	Leu
	X1Hbox4	---	His Phe Asn Arg Tyr Leu Thr Arg Arg Arg Arg Ile Glu Ile Ala His Thr Leu Cys Leu	---	Leu
	X1Hbox5	---	His Phe Asn Arg Tyr Leu Thr Arg Arg Arg Arg Ile Glu Ile Ala Asn Asn Leu Cys Leu	---	Leu
	X1Hbox1	41	Thr Glu Arg Gln Ile Lys Ile Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Lys Glu Ser	60	Asn
	X1Hbox2	---	Thr Glu Arg Gln Ile Lys Ile Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Lys Glu Asn	---	Asn
	X1Hbox3	---	Thr Glu Arg Gln Ile Lys Ile Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Lys Glu His	---	Asn
	X1Hbox4	---	Ser Glu Arg Gln Ile Lys Ile Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Lys Asp Asn	---	Asn
	X1Hbox5	---	Asn Glu Arg Gln Ile Lys Ile Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Lys Asp Thr	---	Asp
	X1Hbox1	61	Asn Leu Ser Ser Thr Leu Pro Gly Gly Thr Gly Ala Ala Ala Asp Ser Leu Ala Glu Gly	80	Glu
	X1Hbox2	---	Lys Ala Ser Ser Pro Ser Ser Asn Ser Gln Glu Lys Pro Glu Thr Glu Glu Glu Glu Glu	---	Glu
	X1Hbox3	---	Lys Glu Glu Ser Asp Gln Thr Pro Asp Ala Gly Glu Glu Ser Thr Ala Pro Thr Thr Thr	---	Glu
	X1Hbox4	---	Lys Leu Lys Ser Met Ser Leu Ala Thr Gly Ser Ser Ala Phe Gln Pro ***	---	***
	X1Hbox5	---	Lys Val Lys Ser Lys Asp Ser Met ***	---	***
	X1Hbox1	81	Lys Glu Glu Lys Gln Glu Asp Ser Glu Gly Gln Gly Lys Glu ***		
	X1Hbox2	---	Glu Glu Glu ***		
	X1Hbox3	---	Ala Asp Asp Asp Lys Asp Lys Glu ***		

8 (XlHbox5) and 34 (XlHbox1) amino acids from the end of the homeodomain (Fig. 2B). The XlHbox1,2,4 and 6 genes have an intron, which is located close to the 5' end of the homeobox (indicated by arrowheads in Figure 2). The same is probably true for XlHbox3, where we have cloned a number of unprocessed transcripts (see Fig. 1B and 2A). The complete XlHbox3 cDNA sequence has been independently obtained and recently published by (19), who call this gene Xhox-36.

It appears that the homeobox region in the frog genes could be extended slightly beyond the 60 amino acid homology. Four out of five genes have a Gly at position -3 (see Fig. 2B) and all have an Asp at position -1. Similarly, all genes have a Lys at position 61 (except for XlHbox1 that has Asn) and a Ser at position 64.

In both the XlHbox1 and XlHbox2 gene there are differences between the cDNA sequences and the previously published genomic sequences. From our studies (unpublished) it is clear that there are two versions of these genes. While the differences in the case of XlHbox1 are minor (Thr at position -2 instead of Ser), in the case of XlHbox2 there are four changes including one in the homeodomain. This is especially surprising, bearing in mind that, for example, Antp and the genomic XlHbox2 only differ by one amino acid in that region. Whether this is of any biological significance remains unclear at the moment.

We have also probed Eco RI digests (Pst I and HindIII, not shown) of genomic Xenopus DNA with our new clones XlHbox3, 4, 5 and 6, as shown in Fig. 3. They all appear to be "single copy" genes, as are XlHbox1 and 2 (3,5). Most genes in Xenopus laevis exist in two versions, frequently giving two bands in genomic Southern blots (XlHbox3, 4, 5 and 6), in keeping with the view that Xenopus laevis is tetraploid in origin (20). We cannot exclude that

Figure 2: Sequence of the five Xenopus homeobox-containing gene.

- A: Nucleotide sequences of all five classes of XlHbox transcripts of the homeobox containing region. Position 1 indicates the start of the homeobox and position 180 is the last nucleotide of the homeobox as defined by comparison with Antp. The arrows indicate the position of synthetic oligonucleotides used for sequencing the homeoboxes. Two blocks of sequences identical in all five genes are highlighted.
 - B: Predicted translation products of nucleotide sequence in panel A. Position 1 indicates the start of the homeo-domain and position 60 the end. The translation stops are indicated by three stars. Amino acids conserved in all five cDNAs are highlighted, and two long stretches of homology are boxed.
- The arrowheads in both panels indicate the predicted position of the intron preceding the homeobox (see text).

Figure 3: Genomic Southern Analysis of X1Hbox3,4,5,6

Xenopus laevis genomic DNA was digested with Eco RI and 8 µg per track were separated on a 0.8% agarose gel, transferred to Nitrocellulose membrane and hybridized to random primed probes (see Mat. and Meth.). For probe fragments see Figure 1A:

A: X1Hbox3:	1 kb	PstI/Eco RI	fragment
B: X1Hbox4:	400 bp	KpnI/Eco RI	fragment
C: X1Hbox5:	500 bp	BamHI/Eco RI	fragment
D: X1Hbox6:	200 bp	SmaI/SmaI	fragment

The arrowheads indicate doublet bands, and size standards are shown.

Xenopus laevis could have truly duplicated copies of most homeobox genes, and experiments addressing this possibility will be attempted in the future.

A Conserved Pentapeptide Found in many Homeobox-containing Genes

All four genes for which we have sequence preceding the homeodomain show a short region of homology the consensus of which is Ile Tyr Pro Trp Met. Figure 4 shows an alignment of this sequence for the 4 X1Hbox-genes, another *Xenopus* gene called Xhox 1A (7), and a similar sequence present in Antp (14,15,16) and cayman β-globin. The Ile Tyr Pro Trp Met pentapeptide, noted by S. Schneuwly to be present in our collection of *Xenopus* cDNAs, has also been reported in a human homeobox containing cDNA by (8) and in three mouse

XlHbox1	Gln	Ile	Tyr	Pro	Trp	Met	Gln	Arg	Met	Asn	Ser	His	Ser	Gly	Val	Gly	Tyr	Gly	Thr	Asp	- HOMEODOMAIN		
XlHbox2	Arg	Ile	Tyr	Pro	Trp	Met	Arg	Ser	Ala	Gly	Ser	Asp	- HOMEODOMAIN										
XlHbox4	Gln	Ile	Phe	Pro	Trp	Met	Arg	Lys	Leu	His	Ile	Asn	His	Asp	Met	Ala	Gly	Pro	Asp	- HOMEODOMAIN			
XlHbox5	Gln	Ile	Tyr	Pro	Trp	Met	Thr	Lys	Leu	His	Met	Ser	His	Glu	Thr	Asp	- HOMEODOMAIN						
Xhox 1A	Val	Val	Tyr	Pro	Trp	Met	Lys	Lys	Ala	His	Ile	Ser	Lys	Ala	Thr	Ser	Thr	Tyr	Ser	Asp	Gly	Glu	- HOMEODOMAIN
Antp	Pro	Leu	Tyr	Pro	Trp	Met	Arg	Ser	Gln	Phe	Gly	Lys	Cys	Gln	Glu	- HOMEODOMAIN							
Globin	Ile	Ile	Tyr	Pro	Trp	Lys	Arg	Arg	Tyr														

Figure 4: Alignment of the consensus Ile Tyr Pro Trp Met in five *Xenopus* homeobox genes, Antp, and cayman β -globin. Identical amino acids are boxed and/or highlighted. Please note the slightly varying distance to the homeodomain, and the extended homology between XlHbox4 and XlHbox5. The sequence of Xhox1A is from the study of Harvey *et al.*, (7).

homeobox containing cDNAs (12). The same consensus sequence occurs in the *Drosophila* genes Deformed (Dfd, 21), Ultrabithorax (Ubx, 22) and to a lesser degree in Caudal (cad, 23) and fushi tarazu (ftz, 24). A more diverged version of this motif is found in all hemoglobins and myoglobin (25) where, from the three dimensional structure derived from X-ray crystallography, it appears to be a bend separating two α -helical regions of the globin protein (26). Since in the homeobox-containing genes it occurs, if present, shortly in front of the homeodomain, one might speculate that it fulfills a similar function in separating the DNA binding moiety (i.e. the homeodomain) from the preceding domain(s) of the protein. The homology in cayman β -globin is the best of all within the globin genes and was picked up in a computer search of the protein sequence database of the protein identification resource of the National Biomedical Research Foundation using the sequence Tyr Pro Trp. It also appears as if the homology between XlHbox4 and XlHbox5 is more extended towards the homeodomain, especially conserving the position of two histidines with respect to the pentapeptide (see Fig. 4).

Expression of Homeobox-containing Genes in Early *Xenopus* Development

In order to examine the pattern of expression for XlHbox3, 4, and 5 we have obtained poly(A)⁺-RNA from different stages of early *Xenopus* development, including oogenesis. While XlHbox1 and 2 (results shown by 3,5) appear to be transcribed at all stages past midblastula, XlHbox3 and 5 expression is confined to the early stages of development (see Fig. 5). The transcription pattern of XlHbox6, not presented in this paper, also peaks during early development, with transcripts being undetectable by the hatching tadpole stage (8). XlHbox3 has a single transcript of 1.7 kb that shows a strong peak of expression at the late neurula stage (17), is not detectable at early neurulation (stage 14) and is only barely detected at the tailbud stage (21), before disappearing entirely. XlHbox5 shows a similar pattern, producing a single transcript of 1.9 kb, that is only expressed at the late neurula and early

Figure 5: Northern blot analysis of X1Hbox3 (panel A), 4 (B) and 5 (C) and actin (D) gene expression during oogenesis and embryogenesis. For staging of oocytes and embryos see Materials and Methods. Two μg of poly-(A)⁺-RNA was loaded in each track. Hybridization was done with nicktranslated insert of each gene at high stringency (see Material and Methods). Size standards are indicated to the left of each panel. The actin probe was a 250 bp fragment of a cytoskeletal actin (clone B9.118).

tailbud stage and is not detectable before and after. X1Hbox4 has a transcript size of 1.5 kb and peaks at the tailbud stage (7). However, it is also present in late neurula and in the swimming tadpole (st 45) at a lower abundance. None of these three genes show multiple transcripts as for X1Hbox1 and 2 (3,5), maternal expression, or expression prior to midblastula in amounts detectable on our Northern blots.

Figure 5D shows a control hybridization with a *Xenopus* actin probe to demonstrate that the RNA preparations of all stages are equivalent. The actin transcription follows exactly the predicted pattern, with a strong expression of two cytoskeletal actin mRNAs in oogenesis. The RNA then decreases in the first stages of embryogenesis and is strongly induced again past the midblas-

Figure 6: SDS-PAGE gel of *in vitro* translation products. Capped SP6 polymerase transcripts of X1Hbox2 (pG52 and pG52 Δ) were translated *in vitro* (see Materials and Methods). 1 μ l of each translation was separated on the gel. The (-) track is a control translation without RNA. The estimated size of the protein products is indicated. Note that deletion of the 5' non-coding leader in pG52 greatly increases translation of X1Hbox2 protein.

tula stage, when the shorter muscle-specific transcript is also turned on (21).

In vitro Translation of X1Hbox2

Comparing the transcript sizes on Northern blots for the X1Hbox genes with the clones obtained, we conclude that X1Hbox1 and X1Hbox2 are near full-length cDNAs (X1Hbox1 = p 10; X1Hbox2 = p 52; see Fig. 1A). We subcloned these inserts into the Gemini 2 plasmid (Promega Biotec) in order to synthesize RNA for *in vitro* translation experiments (28). The results of such an experiment are shown in Figure 6, where two X1Hbox2 clones (pG52, pG52 Δ) were translated in a reticulocyte lysate (Promega Biotec). pG52 Δ has the entire sequence upstream of the Apa I site (see Fig. 1A) deleted, which is untranslated leader sequence (sequence in Wright *et al.*, in press). Although the amount of RNA used in the pG52 and pG52 Δ translation is the same (data not shown), the translation is greatly enhanced when the leader sequence is deleted, yet the protein product is the same length. A similar stimulation of translation was obtained when we deleted the 5' leader of the X1Hbox1 cDNA, which we also observed in microinjected frog oocytes (unpublished data). The reason for this phenomenon was not investigated further, but we think that

Nucleic Acids Research

```

Hox2.1 --- --- --- Met Ser Ser Tyr Phe Val Asn Ser Phe Ser Gly Arg Tyr Pro Asn Gly
Hox1.3 --- --- --- Met Ser Ser Tyr Phe Val Asn Ser Phe Cys Gly Arg Tyr Pro Asn Gly

1 31
X1Hbox4 --- --- --- GGC GGC GGC GGC GGC AAT GTG AGC GGC TCA TAC AGG GAG
Hox2.1 --- --- --- Gly Gly Gly Gly Gly Val Ser Gly Ser Tyr Arg Gly
Hox1.3 Pro Asp Tyr Gln Leu Leu Asn Tyr Gly Ser Gly Ser Ser Leu Ser Gly Ser Tyr Arg Asp
Pro Asp Tyr Gln Leu His Asn Tyr Gly Asp His Ser Ser Val Ser Gly Gln Phe Arg Asn

61 91
X1Hbox4 GCA GGA GGC AAC ATG CAG CCC GGG GGT TAT GGA TCC TAT AAC TAC ACC GGA ATG GAC CTG
Ala Gly Gly Asn Met Gln Pro Gly Ala Tyr Gly Ser Tyr Asn Tyr Thr Gly Met Asp Leu
Hox2.1 Pro Ala Ala --- Met His Thr Gly Ser Tyr Gly --- Tyr Asn Tyr Asn Gly Met Asp Leu
Hox1.3 Ser Ala Ser --- Met His Ser Gly Arg Tyr Gly --- Tyr Gly Tyr Asn Gly Met Asp Leu

121
X1Hbox4 AGT ATC AGC CGG ACT GCA GCT CGC ACC --- --- TAT GGG --- --- GGG GAC AAC TCG ---
Ser Ile Ser Arg Thr Ala Ala Pro Thr --- --- Tyr Gly --- --- Gly Asp Asn Ser ---
Hox2.1 Ser Val Asn Arg Ser Ser Ala Ser Ser Ser His Phe Gly Ala Val Gly Gln Ser Cys Ser Arg
Hox1.3 Ser Val Gly Arg Ser Gly --- --- Ser Gly His Phe Gly Ser --- --- Gly Gln Arg Ala Arg

151 181
X1Hbox4 --- TTC CCA GGC CAG GAG --- AGC AGC AGG TTC AGG --- GCA AAT CAA AAC TCC CCC TTG
Phe Pro Gly Gln Gln --- Ser Ser Arg Phe Arg --- Ala Asn Gln Asn Cys Pro Leu
Hox2.1 Ala Phe Pro Ala Ser Ala Lys Gln Pro Arg Phe Arg Gln Ala Thr Ser Ser Cys Ser Leu
Hox1.3 Ser Tyr Ala Ala Gly Ala Ser Gln Pro Arg Tyr Ser Gln Ala Thr Ser Thr His Ser ---

Ala Ala Pro Ala Pro
211
X1Hbox4 TCG ACC CCC GAC CCC CTT CCT TGT GCC --- --- --- AAA AGC CAG AAA TCT GAG
Ser Thr Pro Asp Pro Leu Pro Cys Ala --- --- --- Lys Ser Gln Lys Ser Gln
Hox2.1 Ser Ser Pro Arg Ser Leu Pro Cys Thr Asn Gly Asp Ser His Gly Ala Lys Pro Ser
Hox1.3 Pro Pro Pro Asp Pro Leu Pro Cys Ser Ala Ser Asp Ser His His Gly Gly Lys Asn Ser

Val Ala Pro Ser Pro Gly
241 271
X1Hbox4 CTG TCG CCC --- --- --- GCG GAT CCT GCC ACC TCC AGC GCC CAC TTG ACT GAG ACA
Gln Gln Thr Ser Ala Ser Ser Gln Thr Asp Gln Ser Thr Pro Arg Ser Gly Ala Thr
Hox2.1 Ala Ser Ser Pro Ser Asp Gln Ala Thr Pro Ala Ser Ser Ala Asn Phe Thr Gln Ile
Hox1.3 Leu Gly Asn Ser Ser Gly --- --- Ala Ser Ala Asn Ala Gly Ser Thr His Ile Ser Ser Arg

301 331
X1Hbox4 GAA GAA ACC AGC GCC TCC TCC GAG ACT GAC GAA TCC ACC CCC AGA AGC GGG GCT CCC ---
Pro Arg Ala Asn Gln Asn Asn Cys Ser Pro Gly Ala Ala Gly Thr Asp Gly Gln Ser
Hox2.1 Gln Gln Ala Ser Ala Ser Ser Gln Thr Pro Gln Gln Ala Ala Ser Gln Leu Ser Ser Pro Ser
Hox1.3 Gln Gly Val Gly Thr Ala Ser Ala Ala Gln Gln Asp Ala Pro Ala Ser Ser Gly Gln

361 391
X1Hbox4 --- CCC GCG GCT CTA CAG GAC AAC TGC TCC GCG GGA GCG GCG ACA GAC GGA CAG AGC
Pro Arg Ala Asn Gln Asn Asn Cys Ser Pro Gly Ala Ala Gly Thr Asp Gly Gln Ser
Hox2.1 Leu Ala Arg Ala Gln Pro Gln Pro Met Ala Thr Ser Thr Ala Pro Gln Gly Gln Thr
Hox1.3 --- Ala Gly Ala Gln Ser Gln Pro --- --- --- Ser Pro Ala Pro Pro Ala Gln ---

421 451
X1Hbox4 CCG CAG ATC TTC CCG TGG ATG AGA AAA CTG CAC ATT AAC CAC GAC ATG GCA GGT CCA GAT
Pro Gln Ile Phe Pro Trp Met Arg Lys Leu His Ile Asn His Asp Met Ala Gly Pro Asp
Hox2.1 Pro Gln Ile Phe Pro Trp Met Arg Lys Leu His Ile Ser His Asn Met Thr Gly Pro Asp
Hox1.3 Pro Gln Ile Tyr Pro Trp Met Arg Lys Leu His Ile Ser His Asn Ile Gly Gly Pro Gln

Ass
HOMEBORX
481 511
X1Hbox4 -GGG AAG AGG GCG AGG ACT GCC TAC ACC CGG TAC CAG ACG CTG GAG CTG GAG AAG GAG TTC-
Gly Lys Arg Ala Arg Thr Ala Tyr Thr Arg Tyr Gln Thr Leu Gln Leu Gln Lys Gln Phe-
Hox2.1 -Gly Lys Arg Ala Arg Thr Ala Tyr Thr Arg Tyr Gln Thr Leu Gln Leu Gln Lys Gln Phe-
Hox1.3 -Gly Lys Arg Ala Arg Thr Ala Tyr Thr Arg Tyr Gln Thr Leu Gln Leu Gln Lys Gln Phe-

541 571
X1Hbox4 -CAT TTC AAT AGG TAC CTG ACC AGG AGG AGG ATC GAG ATC GCC CAC ACC TTG TGC CTC-
His Phe Asn Arg Tyr Leu Thr Arg Arg Arg Arg Ile Gln Ile Ala His Thr Leu Cys Leu-
Hox2.1 -His Phe Asn Arg Tyr Leu Thr Arg Arg Arg Arg Ile Gln Ile Ala His Ala Leu Cys Leu-
Hox1.3 -His Phe Asn Arg Tyr Leu Thr Arg Arg Arg Arg Ile Gln Ile Ala His Ala Leu Cys Leu-

601 631
X1Hbox4 -TCC GAG CGG CAG ATT AAA ATT TGG TTC CAG AAC CGG AGG ATG AAG TGG AAA AAG GAC AAC-
Ser Gln Arg Gln Ile Lys Ile Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Lys Asn Asn-
Hox2.1 -Ser Gln Arg Gln Ile Lys Ile Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Lys Asn Asn-
Hox1.3 -Ser Gln Arg Gln Ile Lys Ile Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Lys Asn Asn-

661 691
X1Hbox4 AAG CTG AAA AGC ATG AGC CTC GCC ACG GCG ACC AGC GCC TTC CAG CCC GAGGGCGTTGCCAG
Lys Leu Lys Ser Met Ser Leu Ala Thr Gly Ser Ser Ala Phe Gln Pro ---
Hox2.1 Lys Leu Lys Ser Met Ser Leu Ala Thr Ala Gly Ser Ala Phe Gln Pro ---
Hox1.3 Lys Leu Lys Ser Met Ser Met Ala Ala Ala Gly Gly Ala Phe Arg Pro ---

721 751 781
X1Hbox4 GACCTGGGCAACTATCGGCAAGTGAAGGCCACAGGACTAGCAGCCACTTCTCATTTCTAGTCGCCCTCCCAATGCTATAA
811 841
X1Hbox4 GTCCCTCATTTGTTTTCGACCAAGTTGCTGCTGCTGCTGCCGGGGAGATCGATGTTTTTCACAGCTCTTAAATGTTAT
871 901
X1Hbox4 TCTAAGCTATTATTAATAAAGAATCATTCGGGGTTTTT

```

there might be sequences in the leader which either destabilize the RNA or inhibit translation. In *Drosophila* many homeotic mRNAs also have long 5' non-coding leaders (14,24,29), which may be subjected to translational regulation.

XlHbox4 Shares Extended Homologies with two Mouse Homeobox containing cDNAs

Figure 7 shows the nucleotide and putative amino acid (aa) sequence of XlHbox4 and a comparison to the translation products of the murine Hox 2.1 (11) and Hox 1.3 cDNAs (12). Judging from the transcript size of XlHbox4 (see Fig. 5, panel B), our cDNA is approximately 400-500 bp short at the 5' end. Figure 4 suggests that we lack the sequence for the first 20-25 amino terminal aa, assuming that the initiator methionine is in a position similar to that of Hox 2.1 and Hox 1.3. The homeodomain of all three genes is identical except for one aa substitution in XlHbox4 (Thr for Ala, at nucleotide pos. 571). The region downstream of the homeodomain shows a high degree of similarity, with the translation stop being in the same position in all three genes. The three genes share regions of similarity preceding the homeodomain (solid boxes in Fig. 7), including the Ile Phe Pro Trp Met motif.

Overall, with the alignment as presented in Fig. 7, XlHbox4 is similar to Hox 2.1 in 166 out of a possible 230 positions, similar being either identical aa or very conservative changes (Thr/Ser, Asn/Gln, Asp/Glu, Val/Ile/Leu, Phe/Tyr). XlHbox4 and Hox 1.3 are conserved in 139 positions, while Hox 2.1 and Hox 1.3 share identical or similar residues in 181 out of a possible 270 positions. Therefore, XlHbox4 is more similar to Hox 2.1 (72%) than to Hox 1.3 (60%), and Hox 2.1 is more homologous to XlHbox4 (72%) than it is to Hox 1.3 (67%). However, the degree of similarity between Hox 2.1 and Hox 1.3 suggests that they have arisen from a common ancestral gene, that may have duplicated (as part of a cluster of homeobox-containing genes) in the mouse genome. Since there is increasing evidence for the existence of such duplicated loci, it becomes very difficult to say whether XlHbox4 represents the true frog homologue of murine Hox 2.1, based purely on sequence comparisons. A detailed functional comparison should give the answer whether such analogies can be drawn. A potential homologue to Hox 2.1 (or Hox 1.3) has

Figure 7: XlHbox4 nucleotide and amino acid sequence. The deduced protein sequence is compared to the mouse genes Hox 2.1 (11) and Hox 1.3 (12). Residues which are homologous (see also text) in at least two sequences are highlighted. The homeobox is indicated with the stippled box. Solid boxes show larger regions of similarity outside of the homeodomain. The polyadenylation signal in XlHbox4 is underlined.

also been cloned from a human source (6,30), and it will be interesting to compare its protein sequence to the three above mentioned genes.

Alignments were made in pairs using the program ALIGN of the National Biomedical Research foundation with the following parameters: mutation data matrix, gap penalty of 6 and 8, and 100 randomizing runs. In Fig. 7, aa were boxed (to indicate regions of similarity) if their value was >0 in the mutation data matrix.

DISCUSSION

We have isolated 36 Xenopus laevis cDNA clones expressed early in development, on the basis of their cross-hybridization with probes containing a homeobox region. Subsequently we divided them into six classes according to their specific homeobox sequence, five of which (XlHbox1 to 5) are described here. The sixth gene, described elsewhere, is potentially a very important and useful tissue and region-specific marker (9). When the putative translation products are compared with each other and to the Drosophila Antp gene, they show a high degree of homology in the homeodomain. Two long stretches of 26 and 17 amino acids are identical in all five genes (Fig. 2B). These regions of identity include the helical regions presumed to be important for DNA binding (discussed by 17,32,33), suggesting that all of these proteins could bind to identical DNA target sequences. Furthermore, we demonstrate that the homology of XlHbox4 to two different mouse (Hox 2.1, Hox 1.3) genes is not restricted to the homeodomain, suggesting that these genes are closely related and must share highly conserved biochemical functions throughout the length of the molecule. This remarkable conservation between mouse (or human, 6,30) and frog homeobox-containing genes is not restricted to XlHbox4; human homologues to XlHbox1 and XlHbox2, with extensive amino acid homologies flanking the homeodomain, have been isolated and described (10,13). Thus, out of six Xenopus homeobox-containing genes that we present in this study, at least three have homologues in man. This suggests that vertebrates share a very related family of homeobox containing genes.

We describe a short region of homology (consensus Ile Tyr Pro Tsp Met, 8), that is found in a number of our Xenopus cDNA clones, Drosophila (14), and mammalian (11) homeobox containing genes. Furthermore, this homology is present to a lesser degree in the globin gene family, where it produces a bend separating two helical regions of the protein, as discussed above. We do not know whether it has the same or a similar function in the homeobox containing genes; however, it is always found close to the 3' end of

the exon preceding the homeobox-exon, at a short but variable distance (10-20 aa) in front of the homeodomain.

It may be possible that homeobox-containing genes will turn out to be of great importance in Xenopus development, similar to the role played in Drosophila embryogenesis. It is clear from our results that some genes (XlHbox3, 5 and 6) are expressed in only a narrow window in embryogenesis, and that other genes (XlHbox1 and 2) have multiple transcripts under differential temporal control (3,5, and our unpublished results). Spatially restricted expression of vertebrate homeobox transcripts has been shown by in situ hybridization. A murine gene, Hox-3, is expressed in the middle part of the spinal cord (34) and, in Xenopus, XlHbox1 is enriched in the cervical spinal cord (35). A sensitive S1 assay has shown that the XlHbox6 gene is expressed in the posterior region of the neural plate (9). The localized expression of transcripts in specific regions along the antero-posterior axis is very reminiscent of the behaviour of Drosophila homeotic genes (17,33,36,37) which are known to control segment identity.

The availability of almost full-length cDNA clones, which can be transcribed and translated in vitro opens experimental possibilities for analyzing further the role of homeoboxes in Xenopus development. Harvey et al. (1986) have already used this method to show that Xenopus homeobox proteins are karyophilic, and we have used it to show that the long 5' non-coding leaders present in homeobox transcripts contain possible translational regulators of the XlHbox2 (Figure 6) and XlHbox1 (unpublished data) mRNAs. With the protein products made in vitro it should be possible to study binding to the promoter regions of various homeobox containing genes which we have isolated from the Xenopus genome (K. Cho et al., unpublished). With the collection of Xenopus homeobox-containing genes we have obtained, including a maternally expressed one (5), it might be possible to study the interactions between them and establish a hierarchy of gene activation (33,38).

MATERIALS AND METHODS

Library Screening:

10⁶ clones of a λ gt10 cDNA library of stage 12 (gastrula) Xenopus embryos (39) were screened with a mixed nick translated probe. The previously described fragments were used: XlHbox 1 (118 bp probe, 3), XlHbox 2 (250 bp probe, 5), and Antp (600 bp probe, 1).

The phage plaques were transferred to nitrocellulose paper (Schleicher & Schull) according to the standard Benton and Davis procedure (40). The hybri-

dization conditions were low stringency (1). After purification the cDNA clones were digested with Eco RI and the insert fragments subcloned into the plasmid vector pAT153.

Sequencing:

For sequencing four oligonucleotides were synthesized with complementarity to three regions within the homeobox, which are most conserved among the previously known sequences of X1Hbox1 and X1Hbox2. Two oligonucleotides complementary to nucleotides 31-51 and nucleotides 136-152, (see Fig. 2A) were synthesized for sequencing towards the 3' end of the homeobox. The two other oligonucleotides extended in the opposite direction and were directed against nucleotides 133-153 and 157-176.

In most cases, heat-denatured pAT153 plasmid subclones served as template, on which a normal dideoxy sequencing procedure was carried out (41). In some cases fragments were cloned into M13 vectors (42). All of the sequences discussed here were determined by independently sequencing both strands.

Northern Blots:

Ovaries, oocytes and embryos were homogenized using a Polytron mixer (PT 10-35) and total RNA was extracted as described in (43). Poly-(A)⁺-RNA was obtained by one passage over an oligo-(dT)-cellulose column (Collaborative Research, Type 3).

Poly-(A)⁺-RNA (2 µg per track, as determined by OD₂₆₀/OD₂₈₀ measurements and a poly-A assay using ³²P-labeled poly-U, 44) was separated on a 1.5% agarose-formaldehyde gel and electrotransferred to Gene-Screen paper (New England Nuclear, see 45). ³²P-nick-translation and hybridization at high stringency conditions were standard procedure (40). An RsaI digest of end-labeled denatured pBR322 DNA was used as molecular weight marker.

Staging of Embryos and Oocytes:

Staging of embryos was according to Nieuwkoop and Faber (39). In Figure 4 pre-mid blastula transition (pre-MBT) corresponds to stage 6 of embryogenesis, late blastula to stage 10, early neurula to stage 14, late neurula to stage 18, tailbud to stage 22-24, hatching to stage 28, and swimming tadpole to stage 40.

Small ovary is taken from an immature frog and contains mostly stage 1 oocytes with connective tissue (staging according to 46). Large oocytes are stage 5 and 6 oocytes which have been manually isolated from an ovary.

Southern Blots:

Xenopus laevis genomic DNA was isolated from red blood cells. Southern

blots were performed according to standard Benton and Davis procedures (40) and transferred to Nitrocellulose membrane (Schleicher & Schull). Hybridization conditions were at high stringency (1), using random primed ³²P-labelled DNA fragments (47).

In vitro transcription and translation:

Insert fragments were subcloned into Gemini 1 or Gemini 2 plasmid (Promega Biotec). Capped RNAs were obtained using capping protocol #2 by Promega Biotec (1986/7 catalogue), with m⁷GpppG (Pharmacia) as cap analogue.

2-3 µg of such RNA were added to a micrococcal nuclease treated rabbit reticulocyte lysate (Promega Biotec) and in vitro translation carried out according to supplied protocol (Promega Biotec). ³⁵S-methionine (New England Nuclear) was used as labelled amino acid.

Subsequently the reaction mixture was separated on a 17.5% SDS-PAGE gel (48) and exposed to Kodak XAR-5 X-ray film for 6 hrs to 2 days.

Acknowledgments:

We wish to thank Georges Martin for excellent technical assistance throughout the work, Richard Harland for providing the cDNA library, Stefan Schneuwly and Walter Gehring for their discussion, and Christopher Wright, Ken Cho and Larry Zipursky for their advice on the manuscript, and John Gurdon for the actin probe. We wish to acknowledge Thomas Burglin for carrying out the computer searches. This work was initiated at the Biozentrum, University of Basel, Switzerland, and supported by grants of the Swiss Nationalfonds and the N.I.H. (grant #HD21502-02).

REFERENCES:

1. McGinnis, W., Levine, M., Hafen, E., Kuroiwa, A. and Gehring, W.J. (1984). *Nature*, **308**, 428-433.
2. Scott, M.P. and Weiner, A.J. (1984). *Proc. Natl. Acad. Sci.*, **81**, 1075-1078.
3. Carrasco, A.E., McGinnis, W., Gehring, W.J. and De Robertis, E.M. (1984). *Cell*, **37**, 409-414.
4. McGinnis, W., Garber, R.L., Wirz, J., Kuroiwa, A. and Gehring, W.J., (1984). *Cell*, **37**, 403-408.
5. Muller, M.M., Carrasco, A.E. and De Robertis, E.M. (1984). *Cell*, **39**, 157-162.
6. Levine, M., Rubin, G.M. and Tjian, R. 1984. *Cell*, **38**, 667-673.
7. Harvey, R.P., Tabin, C.J. and Melton, D.A. (1986). *EMBO J.* **5**, 1237-1244.
8. Mavilio, F., Simeone, A., Giampaolo, A., Faiella, A., Zappavigna, V., Acampora, D., Poiana, G., Russo, G., Peschle, L. and Boncinelli, E. (1986). *Nature*, **324**, 664-668.

9. Sharpe,C.R., Fritz,A., De Robertis,E.M. and Gurdon,J.B. (1987). Cell, in press.
10. Simeone,A., Mavilio,F., Acampora,D., Giampaolo,A., Faiella,A., Zappavigna,V., Esposito,M.D., Pannese,M., Russo,G., Boncinelli,E. and Peschle,C. (1987) Proc. Natl. Acad. Sci. USA, **84**, 4914-4918.
11. Krumlauf, R., Holland, P.W.H., McVey, J.H., and Hogan, B.L.M. (1987). Development **99**, 603-617.
12. Odenwald,W.F., Taylor,C.F., Palmer-Hill,F.J., Friedrich,V., Tani,M. and Lazzarini,R.A. (1987) Genes Dev., **1**(5) 482-496.
13. Meijlink,F, deLaaf,R., Verrijzer,P., Destree,O., Kroezen,V., Hilkens,J. and Deschamps,J. (1987). Nucl. Acid Res., **15**, 6773-6786.
14. Schneuwly,S., Kuroiwa,A., Baumgartner,P. and Gehring,W.J. (1986). EMBO J., **5**, 733-729.
15. Laughon,A., Boulet,A.M., Bermingham,J.R., Laymon,R.A. and Scott,M.P., (1986). Molec. Cell. Biol., **6**, 6476-6481.
16. Stroehrer,V.L., Jorgensen,E.M. and Garber,R.L. (1986). Molec. Cell. Biol., **6**, 4667-4675.
17. Gehring,W.J. (1987). Science, **236**, 1245-1252.
18. Joyner,A.L., Kornberg,T., Coleman,K.G., Cox,D.R., and Martin,G.R. (1985). Cell, **43**, 29-37.
19. Condie,B.G. and Harland,R.M. (1987). Development, **101**, 93-105.
20. Kobel,H.R., and DuPasquier, L. (1986). Genetics of polyploid Xenopus. TIGS, (Dec. 86), 310-315.
21. Regulski,M., Harding,K., Kostriken,R., Karch,F., Levine,M. and McGinnis, W. (1985). Cell, **43**, 71-80.
22. Wilde,L.D. and Akam,M. (1987). EMBO J., **6**, 1393-1401.
23. Mlodzik,M., Fjose,A. and Gehring,W.J. (1985). EMBO J., **4**, 2961-2969.
24. Laughon,A. and Scott,M.P. (1984). Nature, **310**, 25-31.
25. Stryer,L. 1981. Biochemistry, W.H. Freeman and Company, San Francisco.
26. Fermi,G. (1975) J. Mol. Biol., **97**, 237-256.
27. Mohun,T.J., Brennan,S., Dathan,N., Fairman,S. and Gurdon,J.B. (1984). Nature, **311**, 716-721.
28. Krieg,P.A., and Melton,D.A. (1985). EMBO J., **4**, 3463-3471.
29. Mlodzik,M. and Gehring,W.J. (1987). Cell, **48**, 465-478.
30. Hauser,C.A., Joyner,A.L., Klein,R.D., Learned,T.K., Martin,G.R. and Tjian,R. (1985). Cell, **43**, 19-28.
31. Desplan,L, Theis,J. and O'Farrell,P.M. (1985). Nature, **318**, 630-635.
32. De Robertis,E.M., Fritz,A., Goetz,J., Martin,G., Mattaj,I.W., Salo,E., Smith,G.D., Wright,C.V.E., and Zeller,R. (1985) Symp. quant. Biol., **50**, 271-275.
33. Akam,M. (1987). Development, **101**, 1-22.
34. Awgulewitsch,A., Utset,M.F., Hart,C.P., McGinnis,W., and Ruddle,F.H. (1986). Nature, **320**, 328-335.
35. Carrasco,A.E., and Malacinski,G.M. (1987). J. Dev. Biol., in press.
36. Levine,M., Hafen,E., Garber,R.L. and Gehring, W.J. (1983) EMBO J. **2** 2037-2046.
37. Levine,M. and Harding,K. (1987). Oxford Survey on Eukaryotic Genes (in press).
38. Meinhardt,M. (1986). J. Cell Sci. Suppl., **4**, 357-381.
39. Nieuwkoop,P.D. and Faber,J. (1967). Normal Table of Xenopus laevis. North Holland, Amsterdam.
40. Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982). Molecular Cloning, A Laboratory Manual. Cold Spring Harbor Laboratory Press, N.Y.
41. Sanger,F., Nicklen,S. and Coulson,A.R., (1977). Proc. Natl. Acad. Sci.,

- 74, 5463-5467.
42. Messing, J., and Vieira, J. (1982). *Gene*, **19**, 269-276.
43. Feramisco, J.R., Smart, J.E., Burridge, K., Helfman, D.M. and Thomas, G.P. (1982). *J. Biol. Chem.* **257**, 11024-11031.
44. Bishop, J.O., Osbash, M., and Evans, D. (1974). *J. Mol. Biol.* **83**, 75-86.
45. Fritz, A., Parisot, R., Newmeyer, D. and De Robertis, E.M. (1984). *J. Mol. Biol.* **178**, 273-285.
46. Dumont, J.N. (1972). *J. Morph.*, **136**, 153-180.
47. Feinberg, A.P., and Vogelstein, B. (1982). *Anal. Biochem.* **132**, 6-13.
48. Laemmli, U.K. (1970). *Nature*, **227**, 280-285.