Volume 333, number 3, 271–274 FEBS 13137 November 1993

Presence of eight distinct homeobox-containing genes in cnidarians**

Morihiro Naito[a], Hiroshi Ishiguro[a], Toshitaka Fujisawa[b], Yoshikazu Kurosawa[a,*]

[a]*Institute for Comprehensive Medical Science, Fujita Health University, Toyoake, Aichi 470-11, Japan*
[b]*National Institute of Genetics, Yata, Mishima 411, Japan*

Received 28 June 1993; revised version received 2 September 1993

Using the polymerase chain reaction, we identified four different homeobox-containing genes in *Hydra magnipapillata*. Three of them, *cnox1-Hm*, *cnox2-Hm* and *cnox4-Hm*, were equivalent to homeobox genes that had already been identified in other species of cnidarians. *cnox5-Hm* was a new homeobox gene and was very similar to *Mox1* in the mouse. Together with the published data, our results indicate that there are at least eight distinct classes of homeobox genes in cnidarians. These homeobox genes show a maximum of 60 to 77% identity in terms of the amino acid residues in their homeodomains to certain classes of homeobox genes that have been identified in *Drosophila*.

Polymerase chain reaction; Homebox gene; *Mox1*; Evolution

1. INTRODUCTION

The homeobox was first identified as a sequence that is commonly found in several homeotic and segmental genes in *Drosophila melanogaster* [1]. To date, homeobox-containing genes have been isolated from various eukaryotes, including members of the animal, plant and fungal kingdoms [2,3]. In order to identify homeobox genes in various organisms, two different methods have been developed. Since the amino acid sequence of the helix 3 region of homeodomains is highly conserved, degenerate oligonucleotides corresponding to this region function efficiently as probes for the identification of homeobox genes [4]. The polymerase chain reaction (PCR), using two primers that correspond to helices 1 and 3 in the homeodomain, has also been utilized for the identification of homeobox genes [5,6]. Several groups have reported the isolation of homeobox genes from various species of cnidarians [6–10]. In this study, we identified four different homeobox genes in *Hydra magnipapillata*.

2. MATERIALS AND METHODS

DNA was extracted from the whole tissue of *Hydra magnipapillata* by the standard method [11]. PCR was carried out as described elsewhere [12] with the *Hydra* DNA as the template. The two primers used were the same as those described in a previous report [13]. The products of PCR were analyzed by electrophoresis on a 4% agarose gel. The 130-bp fragments were cloned and sequenced. A genomic library was constructed by partial digestion of the *Hydra* DNA with *Sau*3A and ligation of fragments of approximately 17 kb with λdash vectors (Stratagene). The library was screened by the Benton-Davis method [14]. Nucleotide sequences were determined by the chain-termination method [15].

*Corresponding author. Fax: (81) (562) 93–8835.

**The sequences reported in this paper have been deposited in the EMBL database (accession nos. Z22638, Z22639, Z22640).

Abbreviations: PCR, polymerase chain reaction.

3. RESULTS AND DISCUSSION

3.1. *Identification of four different homeobox genes in* Hydra magnipapillata

Genomic DNA of *Hydra magnipapillata* was subjected to amplification by PCR with two primers, which corresponded to the sequences ELEKEF and WFQNRR, respectively [6,13]. The 130-bp fragments among the products of PCR were cloned and sequenced. Fourteen clones were analyzed and were classified as being of four types. The four types were designated *cnox1-Hm*, *cnox2-Hm*, *cnox4-Hm* and *cnox5-Hm*. Using these clones as probes, we screened a genomic library constructed from the *Hydra* DNA, and several positive clones were isolated. Fig. 1A shows the nucleotide sequence of the *cnox1-Hm* gene which clearly encodes a homeodomain. The rule for RNA splicing was used to define the tentative 5′-end of the exon. There is a stop codon 46 bp downstream of the homeobox. Fig. 1B shows the nucleotide sequence of *cnox4-Hm* gene. It also encodes a homeodomain. In the case of the *cnox5-Hm* gene, the N-terminal portion of the homeobox was truncated by cleavage of the restriction site, *Sau*3A, that had been used for cloning (Fig. 1C). Since the complete nucleotide sequence of a gene in *Hydra vulgaris* that is equivalent to *cnox2-Hm* has already been reported [10], no further characterization of *cnox2-Hm* was performed in the present study.

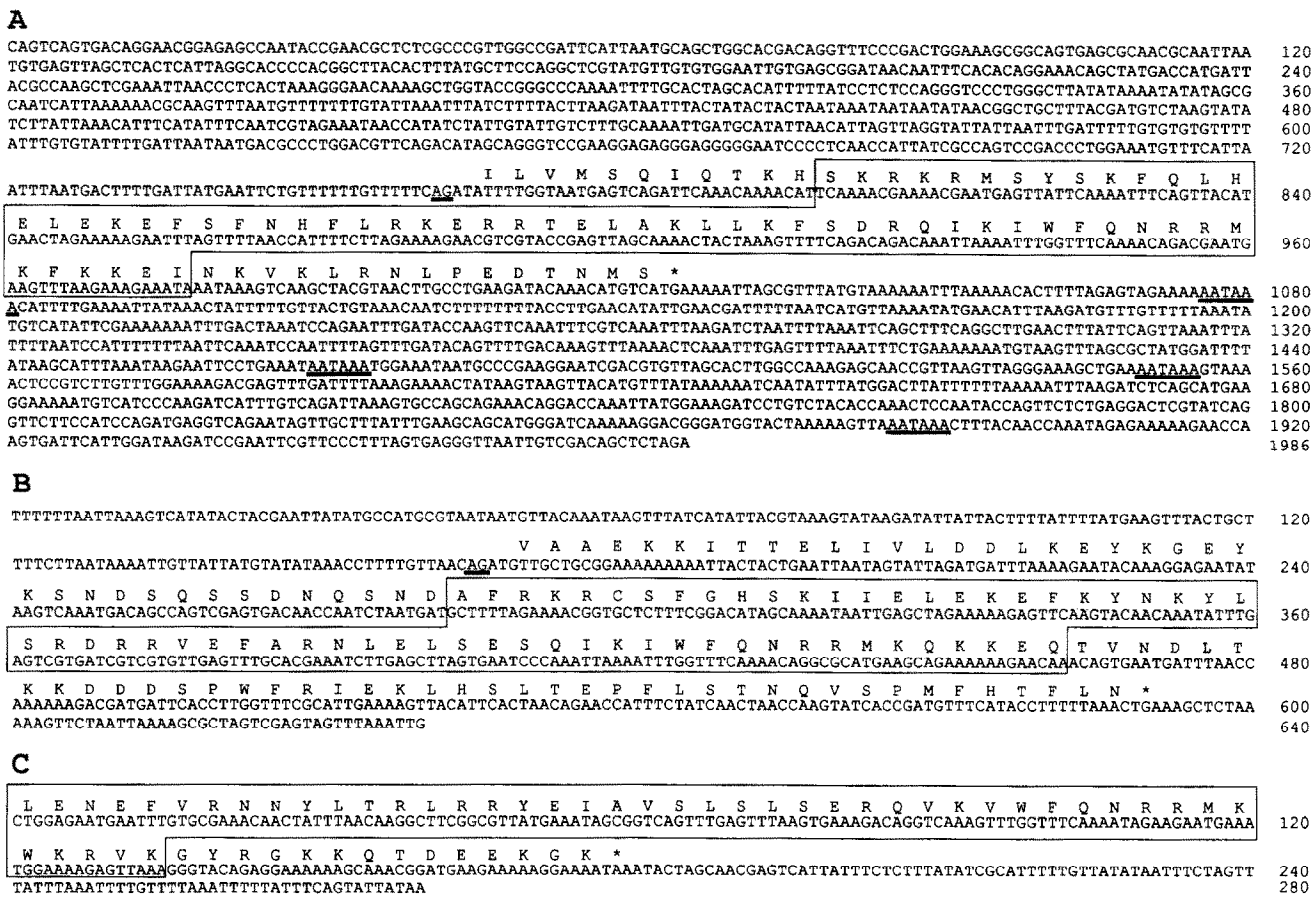

Fig. 1. Nucleotide and deduced amino acid sequences of homeobox genes in *Hydra*. (A) *cnox1-Hm*, (B) *cnox4-Hm* and (c) *cnox5-Hm*. Homeobox regions are boxed. Possible acceptor sites for RNA splicing and possible signals for the addition of polyA are underlined.

3.2. *Classification of homeobox genes isolated from cnidarians*

Homeobox genes have been isolated from several species of cnidarians either by a PCR strategy [6,7,9,10] or by hybridization with degenerate oligonucleotides that correspond to helix 3 of homeodomains [8]. From similarities in amino acid sequence, the nineteen homeobox genes identified to date can be divided into eight distinct classes, as shown in Fig. 2. These sequences were compared with those of the 37 kinds of homeodomain reported to date in *Drosophila* and numbers of identical amino acid residues were counted (Table I). As pointed out by other investigators [8–10], members of class II (*cnox2* in *Hydra vulgaris* and *Chlorohydra viridissima*), class VII (*eveC* in *Acropora formosa*) and class VIII (*msh* in *Chlorohydra viridissima*) were very similar in

Table I

Comparison of amino acid residues in the homeodomains of cnidarians with those of *Drosophila*. The number of amino acid residues that are identical in each case has been determined and is indicated as a percentage. Values above 50% are shadowed. The references for the amino acid sequences deduced from sequences of *Drosophila* homeobox genes are given by Seimiya et al. [19].

	Antp	abdA	ftz	Scr	Ubx	Dfd	lab	pb	AbdB	zen1	zen2	en	inv	eve	prd	BSH4	BSH9	al	bcd	otd	cad	cut	H2.0	ro	Dll	NK-1	NK-2	NK-3	NK-4	msh	ems	BarH2	Om(1D)	bsh	ap	Cf1-a	I-POU
I	58	60	57	60	58	57	50	52	58	55	53	45	45	43	33	32	28	42	33	27	43	30	37	45	42	40	37	43	38	37	42	40	38	45	28	25	28
II	67	62	65	68	60	70	62	68	53	67	65	53	52	52	35	35	33	45	45	37	47	25	40	52	50	57	43	52	43	50	50	47	47	48	28	32	28
III	42	42	37	43	42	42	48	48	47	47	53	50	47	45	37	33	37	38	37	33	37	30	48	45	57	52	45	45	47	47	43	57	55	62	37	25	25
IV	57	57	53	58	57	60	55	52	53	55	55	43	43	50	35	32	32	38	38	32	40	25	35	50	37	45	38	48	42	45	45	43	40	48	30	27	23
VII	52	53	53	55	53	57	55	55	50	55	52	50	47	77	42	43	37	48	40	42	45	32	38	62	43	53	43	48	43	53	52	52	50	53	32	27	28
VIII	43	45	38	45	45	43	48	47	50	45	50	50	50	47	45	42	40	43	42	37	42	28	37	63	57	58	50	52	52	75	50	50	50	53	37	30	30

	10	20	30	40	50	60			
Antp	RKRGRQTYTR	YQTLELEKEF	HFNRYLTRRR	RIEIAHALCL	TERQIKIWFQ	NRRMKWKKEN	Nomenclature	Species	Ref.
I	S--K-MS-SK	F-LH------	S--HF-RKE-	-T-L-KL-KF	SD--------	-----F---I	cnox1-Hm	*Hydra magnipapillata*	*
			LSHF-KKE-	-VDL-KQ-N-	S-----		cnox1	*Eleutheria dichotoma*	(7)
			SCHF-KKE-	---L-KQ-S-	------		SAox3	*Sarsia sp.*	(6)
			--HF-KKE-	-T-LSKK-N-	S-----		cnox1	*Hydra vulgaris*	(10)
			--HF-KKE-	-A-L-KH-N-	S-----		cnox3	*Hydra vulgaris*	(10)
II	S--I-TA--S	I-L-------	QN----S-L-	--Q--AI-D-	--K-V-----	---V----DK	cnox2	*Hydra vulgaris*	(10)
			QN----S-L-	--Q--AI-D-	--K-V--		cnox2-Hm	*Hydra magnipapillata*	*
			N----S-L-	--Q--AM-D-	--K-V-		cnox2-Ed	*Eleutheria dichotoma*	(7)
			N----S-L-	--Q--AM-D-	--K-V-		cnox2-Hs	*Hydractinia symbiolongicarpus*	(7)
	S--I-TA--S	I-L-------	QN----S-L-	--Q--AI-D-	--K-V-----	---V----DK	cnox2	*Chlorohydra viridissima*	(8)
			N----S-L-	--Q--AM-D-	--K-V-		SAox2	*Sarsia sp.*	(6)
III	CRKP-TVFSD	L-LMV--R--	NNRK--STPQ	-TNL-DR-G-	NQT-V-T-Y-	---------T	cnox3	*Chlorohydra viridissima*	(8)
				-M-	-DN-V-T-Y-	-----L-RHI	cnox4	*Chlorohydra viridissima*	(8)
IV	AF-K-CSFGH	SKII------	KY-K--S-D-	-V-F-RN-E-	S-S-------	-----Q---Q	cnox4-Hm	*Hydra magnipapillata*	*
	CSFGH	RKII---R--	KY------D-	-L-F-RN-D-	S-S---V---	-----Q---Q	cnox1	*Chlorohydra viridissima*	(8)
V		--N--	VR-N----L-	-Y---VS-S-	S---V-V---	-------RVK	cnox5-Hm	*Hydra magnipapillata*	*
VI			--K----A-	-V---QI-K-	--S---		SAox1	*Sarsia sp.*	(6)
VII	TR-Y-TAF--	E-LSR-----	LREN-VS-T-	-S-L-SM-N-	S-TT------	-----A-RRR	eveC	*Acropora formosa*	(9)
VIII	NRKP-TPFSV	N-L-T--QK-	KRKQ--SISE	-A-LSEL-R-	--T-------	---A-Q-RSK	msh	*Chlorohydra viridissima*	(8)

Fig. 2. Classification of homeobox genes in cnidarians. Fifteen homeobox genes from various species of cnidarians have been reported [6–10]. We identified four different homeobox genes in *Hydra magnipapillata*. From similarities among sequences, they have been divided into eight classes. The amino acid sequence encoded by *Antp* gene of *Drosophila* is used as a reference sequence. Bars indicate amino acids that are the same as those encoded by *Antp*. The nomenclature for the homeobox genes has not yet been unified. Therefore, they are tentatively designated as members of classes I to VIII.

terms of amino acid residues to *Dfd*, *eve* and *msh* in *Drosophila*, respectively. The member of class V (*cnox5-Hm*) was 87% identical in a region of 45 amino acid residues to *Mox1* in mouse [16] but the equivalent gene has not yet been reported in *Drosophila*. Members of class I (*cnox1-Hm*) and class IV (*cnox4-Hm*) were rather similar to homeobox genes at HOM loci in *Drosophila*. Although *cnox3* and *cnox4* in *Chlorohydra viridissima* were placed in different classes by Schummer et al. [8], the sequence in helix 3, TWYQNRR, is characteristic of the *BarH2* and *Om(1D)* classes in *Drosophila* [17,18]. Therefore, *cnox3* and *cnox4* are placed in the same class in Fig. 2. *SAox1* in *Sarsia sp.* seems to be different from the other homeobox genes. Thus, we divided the nineteen homeobox genes into eight classes as shown in Fig. 2.

It is quite likely that there are more homeobox genes in cnidarians. For instance, Seimiya et al. [19] identified two homeobox genes, *prox1* and *prox2*, in the most primitive metazoan, the sponge (*Ephydatis fluviatilis*), and they showed that the amino acid sequences of the homeodomains encoded by *prox1* and *prox2* were 72% and 62% identical to those encoded by the *NK-3* and *Om (1D)* genes of *Drosophila*, respectively. They concluded that, when the metazoa appeared during the course of evolution, the multiple and distinct classes of homeobox genes that have been identified in higher organisms already existed. Indeed high percentages of identical amino acid residues are found when specific combinations of homeobox genes are examined, such as when members of class II, class VII and class VIII of cnidarians are compared with *Dfd*, *eve*, and *msh* of *Drosophila*, respectively, as well as when class V of cnidarians is compared with *Mox1* of the mouse. These results are consistent with the above conclusion. The relatively low degree of similarity between some homeobox genes of cnidarians and those of *Drosophila* for example, classes I, III and IV, suggests that there may still be homeobox genes that remain to be identified in cnidarians and even in *Drosophila*.

Acknowledgements: We thank Drs. Y. Takagi, I. Ishiguro, and K. Fujita for their encouragement, and Ms. A. Nagata for preparation of the manuscript. This work was supported in part by grants from Japan Foundation for Aging and Health; Ministry of Education, Science and Culture in Japan; Fujita Health University.

REFERENCES

[1] McGinnis, W., Levine, M.S., Hafen, E., Kuroiwa, A. and Gehring, W.J. (1984) Nature 308, 428–433.
[2] Scott, M.P., Tamkun, J.W. and Hartzell III, G.W. (1989) Biochem. Biophys. Acta 989, 25–48.
[3] Vollbrecht, E., Veit, B., Sinha, N. and Hake, S. (1991) Nature 350, 241–243.
[4] Bürglin, T.R., Finney, M., Coulson, A. and Ruvkum, G. (1989) Nature 341, 239–243.
[5] Kamb, A., Weir, M., Rudy, B., Varmus, H. and Kenyon, C. (1989) Proc. Natl. Acad. Sci. USA 86, 4372–4376.
[6] Murtha, M.T., Leckman, J.F. and Ruddle, F.H. (1991) Proc. Natl. Acad. Sci. USA 88, 10711–10715.
[7] Schierwater, B., Murtha, M., Dick, M., Ruddle, F.H. and Buss, L.W. (1991) J. Exp. Zool. 260, 413–416.
[8] Schummer, M., Scheurlen, I., Schaller, C. and Galliot, B. (1992) EMBO J. 11, 1815–1823.
[9] Milles, A. and Miller, D.J. (1992) Proc. R. Soc. Lond. B 248, 159–161.
[10] Shenk, M.A., Bode, H.R. and Steele, R.E. (1993) Development 117, 657–667.

[11] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[12] Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) Science 239, 487–491.
[13] Naito, M., Kohara, Y. and Kurosawa, Y. (1992) Nucleic Acids Res. 20, 2967–2969.
[14] Benton, W.D. and Davis, R.W. (1977) Science 196, 180–182.
[15] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[16] Candia, A.F., Hu, J., Crosby, J., Lalley, P.A., Noden, D., Nadeau, J.H. and Wright, C.V.E. (1992) Development 116, 1123–1136.
[17] Higashijima, S., Kojima, T., Michiue, T., Ishimaru, S., Emori, Y. and Saigo, K. (1992) Genes Dev. 6, 50–60.
[18] Tanda, S. and Corces, V.G. (1991) EMBO J. 10, 407–417.
[19] Scimiya, M., Ishiguro, H., Miura, K., Watanabe, Y. and Kurosawa, Y. (1993) (submitted).