The EMBO Journal vol.11 no.5 pp.1815-1823, 1992

HOM/HOX homeobox genes are present in hydra (*Chlorohydra viridissima*) and are differentially expressed during regeneration

Michèl Schummer, Imke Scheurlen[1], Chica Schaller[1] and Brigitte Galliot[2]

Zentrum für Molekulare Biologie, Im Neuenheimer Feld 282, 6900 Heidelberg, FRG
[1]Present address: Center for Molecular Neurobiology, Martinistrasse 52, 2000 Hamburg 20, FRG
[2]Corresponding author

Communicated by C.Schaller

Hydra, a diblastic animal consisting of two cell layers, ectoderm and endoderm, is one of the most ancient animals displaying an anteroposterior axis with a head and a foot developing from an uncommitted gastric region. As such, hydra is an interesting model for studying the presence and function of homeobox genes in a phylogenetically old organism. By screening a *Chlorohydra viridissima* cDNA library with a 'guessmer' oligonucleotide, we have cloned several such cnidarian homeobox-containing genes (*cnox* genes). Two of these, *cnox1* and *cnox2*, display *labial* and *Deformed* type homeodomains respectively and could represent two ancestral genes of the HOM/HOX complexes; *cnox3* exhibits some similarity to the *BarH1* and the *distal-less* type homeodomains and a fourth gene is highly related to the *msh/Hox7* type of homeodomain. We used quantitative PCR to study levels of expression of these genes along the body axis and during head regeneration. In all cases, the expression in heads was stronger than that in the gastric region. *cnox1* transcripts dramatically peaked within the first hours of head regeneration, whereas *cnox2* and *cnox3* reached their maximal levels 1 and 2 days after cutting respectively. This differential expression of homeobox genes at various stages of regeneration suggests that they play specific roles in regenerative processes.

Key words: Deformed/hydra/*labial*/*msh*/quantitative PCR

Introduction

The 61 amino acid homeodomain encoded by the homeobox is able to recognize and specifically bind DNA (Gehring *et al.*, 1990; Wolberger *et al.*, 1991) and homeoproteins have been shown to modulate gene expression. Most of the homeobox genes present in arthropods and in vertebrates are developmentally regulated. So far the most primitive organisms in which they have been detected are nematodes (Bürglin *et al.*, 1989), planaria (Garcia-Fernandez *et al.*, 1991), echinoderms (Dolecki *et al.*, 1986) and cnidarians (Murtha *et al.*, 1991; Schierwater *et al.*, 1991). Homeotic selector genes which specify parasegmental identity in *Drosophila* are organized in two complexes, the Antennapedia and Bithorax complexes. In other arthropods such as the beetle *Tribolium*, a unique HOM complex has been described (Beeman, 1987). In *Drosophila* these genes comply to the principle of spatial colinearity, i.e. their expression along the anteroposterior axis follows the same order as the arrangement of the genes along the chromosome (Lewis, 1978; Kaufman *et al.*, 1980). Vertebrate homologues of HOM genes are also organized in clusters (HOX clusters) and expressed according to this rule (Duboule and Dollé, 1989; Graham *et al.*, 1989); moreover a HOM/HOX-like complex has recently been identified in *Caenorhabditis elegans* (Bürglin *et al.*, 1991). From this it is concluded that homeobox genes located within HOM/HOX clusters could in all organisms play a role in the specification of cell fate according to the anteroposterior position. If so, the ancestral HOM/HOX cluster should have been present early in evolution when nematodes, arthropods and chordates diverged and it might have been responsible for a 'conserved molecular representation of front, middle and back' in the ancestral organism (Akam, 1989). Since hydra is phylogenetically very close to a common, diblastic ancestor of both arthropods and vertebrates, and since it is one of the oldest animals which displays an anteroposterior pattern of differentiation, it is therefore a good candidate to be investigated for the presence and the regulation of genes representing the HOM/HOX genes' ancestors. In addition, hydra is able to regenerate its head or foot from the gastric region. These regenerative processes represent an anterior or posterior development from an undifferentiated region, thus mimicking embryonic development.

Here, we present evidence for the existence of five hydra homeobox genes cloned from a multiheaded mutant of *C.viridissima*. Two of them are closely related to HOM/HOX genes. We used a PCR strategy to study quantitative changes in their expression along the body axis and during head regeneration. Differential expression detected at various stages of head regeneration suggests that representative genes of the HOM/HOX ancestors have specific roles in differentiation along the anteroposterior axis of hydra.

Results

Cloning of hydra homeobox-containing genes

We used a 'guessmer' oligonucleotide derived from the most conserved part of the *Antennapedia* (*Antp*) homeodomain (Figure 1), following the hydra codon usage as deduced from various hydra genes [STK (Bosch *et al.*, 1989), actin (Fisher and Bode, 1989) and cyclin B (I.Scheurlen, unpublished)]. We isolated five homeobox-containing clones from a *C.viridissima* cDNA library: *cnox1*, *cnox2*, *cnox3*, *cnox4* and *msh*. The designation *cnox* stands for cnidarian homeoboxes. Comparison of these homeodomains with that of *Antp* shows that the four invariant positions in the third α-helix (WF-N-R) are conserved except for the substitution of the Phe50 residue by Tyr50 in the *cnox3* and *cnox4* sequences (see below); the eight highly conserved positions amongst arthropod and vertebrate homeodomains are also

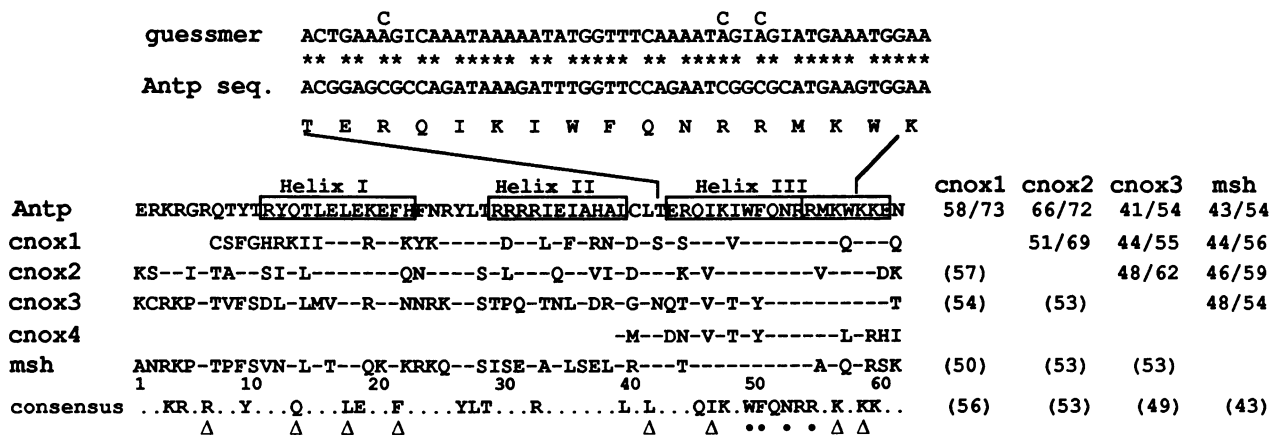

Fig. 1. Strategy for cloning and sequence of five homeobox genes from hydra. The guessmer oligonucleotide used for screening is aligned with the corresponding *Antp* nucleotide and peptide sequence. Five different hydra homeobox genes, *cnox1*, *cnox2*, *cnox3*, *cnox4* and *msh*, were found. Their sequence is compared with the *Antp* homeodomain. Identical amino acids are represented by hyphens. A consensus sequence corresponding to amino acids identical between the hydra homeodomains and *Antp* has been derived. Numbers on the right give the levels of similarity in percent between different genes. The number giving the DNA similarity is in brackets; the last line compares hydra homeoboxes with *Antp*. The peptide relatedness is expressed by two percentages, the first is the percentage of identical amino acids, the second includes conservative substitutions (Val ⟷ Leu ⟷ Ile, Thr ⟷ Ser, Gln ⟷ Asn, Glu ⟷ Asp, Lys ⟷ Arg, Phe ⟷ Tyr). The black points below the consensus sequence indicate the four invariant positions amongst homeodomains, the triangles mark amino acids at the eight highly conserved positions.

conserved (Figure 1). These sequences contain the critical residues necessary for the helix−turn−helix secondary structure required for DNA binding (Scott *et al.*, 1989; Gehring *et al.*, 1990; Wolberger *et al.*, 1991).

Of these five hydra homeodomains, two, *cnox1* and *cnox2*, are related to *Antp* (or class I homeodomains); these two genes are 58% and 66% identical to *Antp* respectively (73% and 72% similar when conservative substitutions are considered). *cnox3* and *msh* are < 50% identical and do not belong to this class (Figure 1). These homeodomains are not closely related to each other; *cnox1* and *cnox2* are only 51% identical (69% similar), while the others are more divergent (Figure 1).

The homeodomains that are most closely related to *cnox1* are those from the chicken *Chox1* (62%), *Drosophila Deformed* (*Dfd*, 62%), human *HOX3E* (60%) and mouse *Hox-2.5* (60%) homeoproteins (Figure 2a). However, when conservative changes are taken into account, the chicken *Ghox-lab*, mouse *Hox-2.9*, human *HOX2I* and *Drosophila Dfd* are, with 75% similarity, the most closely related homeodomains to *cnox1*. *Chox1*, *Ghox-lab*, *Hox-2.9* and *HOX2I* belong to the *labial* (*lab*) family, whereas *Dfd* and *HOX3E* are members of the *Dfd* sub-family and *Hox-2.5* is a mouse *Abdominal B* (*AbdB*) cognate (see legend of Figure 2 for references). Comparing amino acids at different positions where they are variable between different homeodomain classes, but constant for a given family, we observe that the Gln57, which is specific to the *lab* family, is indeed present in the *cnox1* sequence. A similar conservation is seen for Phe9, and for the Thr44 which is substituted by a conservative serine residue in *cnox1*. In contrast, residues which are specific for the *Dfd* or *AbdB* families are largely not found in *cnox1*.

The allocation of *cnox2* to a particular class is more obvious. Indeed, the *cnox2* homeodomain is 70% identical with that of *Hox-1.4*, 69% with that of *Hox-1.3*, *Chox-1.4* and *Dfd* (Figure 2b). *Hox-1.4*, *Chox-1.4* and *Dfd* are all members of the same *Dfd*-like family whereas *Hox-1.3* is closely related to it (Scott *et al.*, 1989). Including conservative substitutions *cnox2* is 77% similar, regardless of species, to the *Dfd/Hox-1.4* family (Figure 2b). Therefore *cnox2* can be considered as an ancestral *Dfd* type homeodomain.

The *cnox3* homeodomain displays a particular feature: the Phe50 in the third α-helix is replaced by a Tyr. This Phe50 residue is one of the four positions invariant in all metazoan homeodomains examined so far with the exception of that of *Drosophila BarH1* (Kojima *et al.*, 1991). This 'atypical' Tyr50 residue is also found in another hydra homeodomain, *cnox4*, the sequence of which is too incomplete to allow further classification (Figure 2c). *cnox3* has the highest level of identity with *BarH1* and *Distal-less* (*Dll*) type of homeodomains (57% identity in both cases), as well as the highest levels of similarity when conservative substitutions are considered (64 and 63%, respectively).

The fifth hydra homeodomain we examined is highly similar to the *Drosophila msh* (muscle specific homeobox) homeodomain and to its vertebrate counterparts mouse *Hox-7*, quail *Quox-7* and chicken *Chox-7* (Figure 2d). The identity at the DNA level is 59% with *msh*, 61% with *Hox-7*; that at the protein level is higher, 74% with *msh* and 72% with *Hox-7*. This homology extends to the first amino acids flanking the homeodomain and reaches 82% with the *msh* homeodomain when conservative changes of amino acids are taken into account. Therefore this hydra homeobox gene probably represents the ancestor of the *msh/Hox7* gene family and has been named *msh*.

Outside of the homeodomain we searched for the hexapeptide sequence (VYPWMK) that is usually located ~15 amino acids in front of *Drosophila* or vertebrate homeodomains (Mavilio *et al.*, 1986). In *cnox2* and *cnox3* sequences this motif is partially conserved (LYP) and located 43 and 42 amino acids upstream of their respective homeodomains. In both cases the additional stretch of 21 amino acids separating this conserved LYP motif from the homeodomain is rich in Glu, Gln or Asn residues (Figures 3 and 4). In the case of *cnox2*, surprisingly, the region surrounding this LYP motif displays some conservation with *Dfd* type homeoproteins (~35% over 40 amino acids, Figure 4). Additional homologies are also detected in the

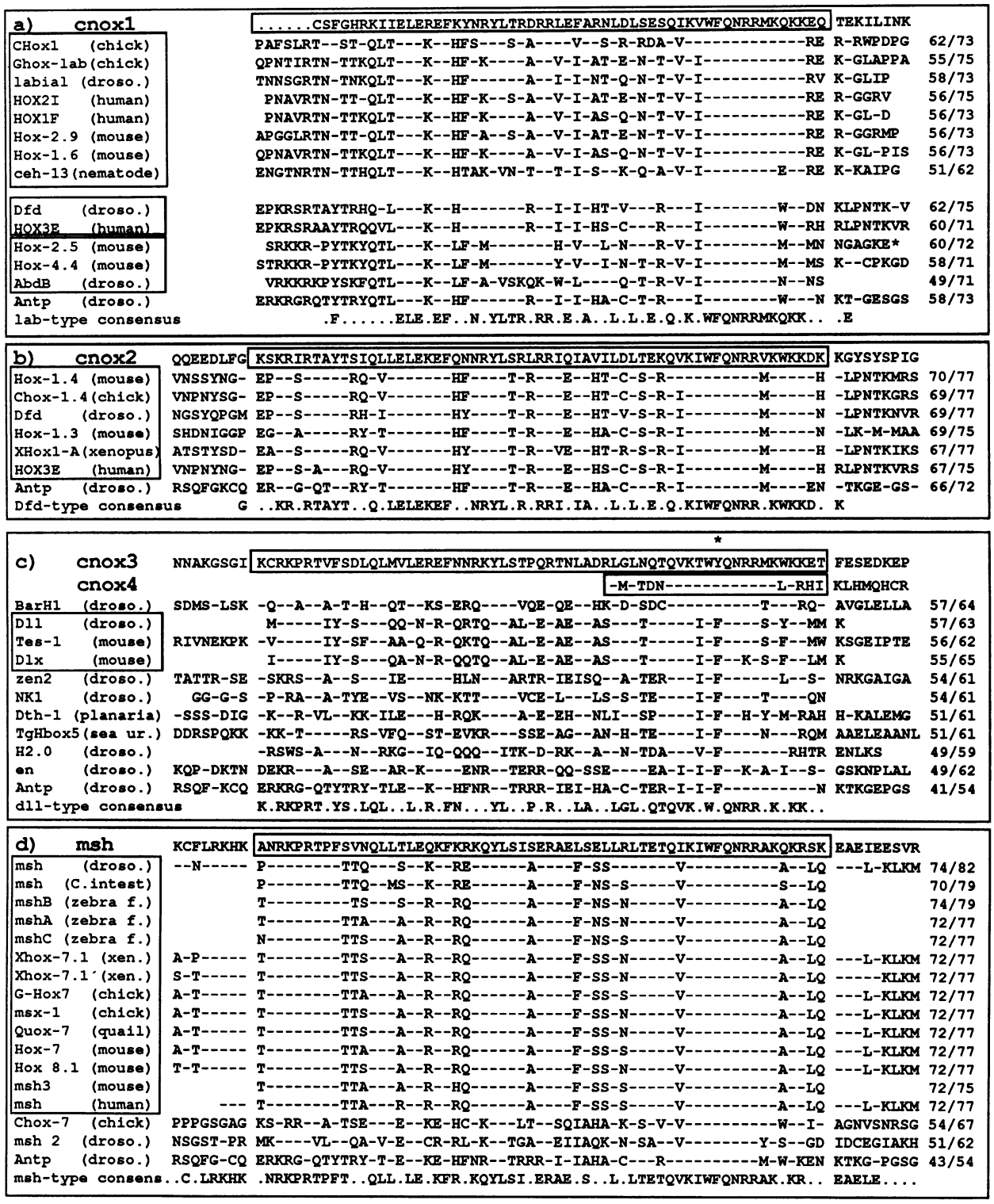

Fig. 2. Alignment of *C.viridissima* homeodomains with the most closely related *Drosophila* or vertebrate homeodomains: The amino acid sequence of hydra homeodomains have been boxed. Hyphens correspond to identical amino acids. Numbers on the right give the peptide similarity as in Figure 1. The names of homeodomains belonging to the same family are boxed. A consensus sequence corresponding to amino acids identical to those of the most related homeodomain family (or conservatively substituted) is indicated for each hydra homeodomain (except *cnox4*). The asterisk above the *cnox3* sequence indicates the unusual Tyr residue. The references to the sequences of *Antennapedia* (*Antp*), *labial* (*lab*), *Hox-1.6*, *Deformed* (*Dfd*), *Abdominal B* (*AbdB*), *Hox-1.4*, *Hox-1.3*, *Xhox1-A*, *zerknüllt 2* (*zen2*), *H2.0* and *engrailed* (*en*) are in Scott *et al.* (1989). The human homeodomain sequences (*HOX2I*, *HOX1F* and *HOX3E*) are in Boncinelli *et al.* (1991). References to other homeodomains are as follows: *CHox1*: Rangini *et al.*, 1989; *Ghox-lab*: Sundin *et al.*, 1990; *Hox-2.9*: Frohman *et al.*, 1990; *ceh-13*: Schaller *et al.*, 1990; *Hox-2.5*: Bogarad *et al.*, 1989; *Hox-4.4*: Dollé and Duboule, 1989; *Chox-1.4*: Sasaki *et al.*, 1990; *BarH1*: Kojima *et al.*, 1991; *Distal-less* (*Dll*): Cohen *et al.*, 1989; *Tes-1*: Porteus *et al.*, 1991; *Dlx*: Price *et al.*, 1991; *NK1*: Kim and Nirenberg, 1989; *Dth-1*: Garcia-Fernandez *et al.*, 1991; *TgHbox5*: Wang *et al.*, 1990; *msh* (*Drosophila*): Gehring, 1987; Robert *et al.*, 1989; *msh* (*Ciona intestinalis*), *mshA*, *mshB*, *mshC* (zebra fish), *msh3*: Holland, 1991; *Xhox-7.1* and *Xhox-7.1'* (*Xenopus laevis*): Su *et al.*, 1991; *G-Hox 7*, *msh* (human): Suzuki *et al.*, 1991; *Msx-1*: Yokouchi *et al.*, 1991; *Hox-7*: Robert *et al.*, 1989; Hill *et al.*, 1989; *Hox 8.1* (= *msh2*): Monaghan *et al.*, 1991; Holland, 1991; *Quox-7*: Takahashi and Le Douarin, 1990; *Chox-7*: Fainsod and Gruenbaum, 1989; *msh2*: Bodmer *et al.*, 1990.

cnox1

```
GGTGTTCTTTCGGGCATAGAAAAATAATTGAATTAGAGAGAGAATTTAAGTATAACAGGTATTTA  65
  C  S  F  G  H  R  K  I  I  E  L  E  R  E  F  K  Y  N  R  Y  L     21
TCACGTGATCGTCGTCTTGAATTTGCACGAAATCTTGACCTTAGTGAGTCCCAGATTAAAGTTTGG 131
T  R  D  R  R  L  E  F  A  R  N  L  D  L  S  E  S  Q  I  K  V  W     43
TTTCAAAATCGACGCATGAAACAAAAAAAAGAGCAGACGGAGAAAATTCTAATAAACAAAAAGTTT 197
F  Q  N  R  R  M  K  Q  K  K  E  Q  T  E  K  I  L  I  N  K  K  F     65
AAAGCAAATTGGCTTCATAATGAAAATTCACGTTCAATAAAACGACAGTGTTTAACAAGTGGACAA 263
K  A  N  W  L  H  N  E  N  S  R  S  I  K  R  Q  C  L  T  S  G  Q     87
GAATCCCCTATATTTCCTCCTCTTTTGAACTAAGAGCTTTAACTTAAAAAGCACCGATAAAGAAGA 329
E  S  P  I  F  P  P  L  L  N  *                                      97
GACTAAATCAAAACGAATCATTACGATTGGGTCTGGCTCCTTTCTTTATTTCACTGGTTATCTTGT.395
TATGTGATAATAAGCTTAGTTGAGACCTTTTTTACCTGATTTTTATCTGACCTGAAACCATATCAA 461
CACTTCTGCTAGAAAAAAATATTTAATTCAAATGATTTAATCAAGTTTGGATCGCGAATTTAGCTTT 527
CGTTAACTTATGAAAACATTCCCTTGAACTTTTTAACCCGGCGGTCATTTACTAAGTAACCTCTTG 593
AAGAAGAATTTAGTTGTGGCTAAAGTAATTGCTTAAAATACCAATTTTCGTCAACGGTTTTGATAT 659
GATAAAGCCCATGTATGGTGAAGAAGCACACTTCGGTGAATGCTGATGTCGTTAGAGTAGGCTTCA 725
AGTAAAAACCATGCTAATATTGAAATTAAAATATTTACACGAGAGAACAATAACAATATCAAATTT 791
TTAGATTTGTAGAAATTTTCATAAATGAAGTTTTTA                               827
```

cnox2

```
TTGATAGACTCGATAATACATGAAAAAGAGAAGTTTAAGATGCGGCAACAACCAGCAACAGCTTTT  66
L  I  D  S  I  I  H  E  K  E  K  F  K  M  R  Q  Q  P  A  T  A  F     22
CTATTTCGTGAGTCGTCCCCTTCAGACCGATCCCCGAGTTATTCACCGGTTGCCTCAATGATTAGA 132
L  F  R  E  S  S  P  S  D  R  S  P  S  Y  S  P  V  A  S  M  I  R     44
TATTCTCCTTCATCTTCTCCAAGAAGTTTAGATTCACCTGTTAATCAATTGGAAAGACATCCTCTA 198
Y  S  P  S  S  S  P  R  S  L  D  S  P  V  N  Q  L  E  R  H  P  L     66
GAAAGAGTGCACCAAGTAATCAGTTGTGTAAGAGGACCTTCTATGTGTAATTGCTGTCGGCCGCAT 264
E  R  V  H  Q  V  I  S  C  V  R  G  P  S  M  C  N  C  C  R  P  H     88
CCCGTTCAGCCCATATGCACCGTATGCGAATCTAGGGAGCCGGGCGAAGGGGCTTCTTCACAATAT 330
P  V  Q  P  I  C  T  V  C  E  S  R  E  P  G  E  G  A  S  S  Q  Y    110
CCATACACCAGAGAACCACATGAGCATTCTCCTAGAGATTTATATGGAATTGAAAGGTCGAGACTT 396
P  Y  T  R  E  P  H  E  H  S  P  R  D  L  Y  G  I  E  R  S  R  L    132
TATCCTATTACATCTCCTTTGCACGGACAAAGAGCACAGTTTCCCCAAACTATGTTTACGACCTA 462
Y  P  I  T  S  P  L  H  G  Q  R  A  Q  F  S  P  N  Y  V  Y  D  L    154
GAAGTACGCCATACTCGTCAACTTCAATTGCAACACCAGCAACAAGAAGAGGACCTTTTCGGGAAA 528
E  V  R  H  T  R  Q  L  Q  L  Q  H  Q  Q  Q  E  E  D  L  F  G  K    176
TCAAAACGGATCCGAACGGCTTACACGAGCATACAACTGCTTGAACTTGAAAAAGAATTTCAAAAC 594
S  K  R  I  R  T  A  Y  T  S  I  Q  L  L  E  L  E  K  E  F  Q  N    198
AATCGATATTTATCGAGACTAAGGAGAATCCAGATCGCTGCAATTTTGGATTTAACTGAAAAACAA 660
N  R  Y  L  S  R  L  R  R  I  Q  I  A  A  I  L  D  L  T  E  K  Q    220
GTAAAAATATGGTTTCAAAACAGACGTGTTAAATGGAAAAAAGATAAGAAAGGATATAGCTATTCC 726
V  K  I  W  F  Q  N  R  R  V  K  W  K  K  D  K  K  G  Y  S  Y  S    244
CCCATTGGCAGCCCTAGCTCTCCAGAGTAATTTCTGATATCACGTTTTTCTGTCTTTCCACAAGAA 792
P  I  G  S  P  S  S  P  E  *                                        253
ATTATTTTCTTATACAATCCTCTTTCGAACTTTCAAAGATTTTATAAAGAAAACCAGGAGAAAAAT 858
AAAAGAAGGAAAATAAAATGGCAAGAGTTCCATGGTTAATTGAAAAATGTACATAAAAAATATTTA 924
ATAATTCATGAATTTTTGCGATTTAAGCTTTTTGTATTTTGATTT                      970
```

cnox3

```
AATTTATATCCAATATTAAATACTGATCAAAATCACTGTACGTACGCCAAAGAAGATTCTCTGATA  66
N  L  Y  P  I  L  N  T  D  Q  N  H  C  T  Y  A  K  E  D  S  L  I    22
CCTGAAGTAGAAGAGCCATCTACATATTTGCAGTTGAAACAAAATAATGCTAAAGGCAGCGGAATA 132
P  E  V  E  E  P  S  T  Y  L  Q  L  K  Q  N  N  A  K  G  S  G  I    44
AAATGCAGGAAACCTCGGACGGTATTTTCTGATCTTCAACTCATGGTGCTTGAACGTGAATTCAAC 198
K  C  R  K  P  R  T  V  F  S  D  L  Q  L  M  V  L  E  R  E  F  N    66
AATCGCAAATATTTATCAACGCCTCAGCGGACAAACTTAGCAGACCGACTTGGCCTGAATCAGACA 264
N  R  K  Y  L  S  T  P  Q  R  T  N  L  A  D  R  L  G  L  N  Q  T    88
CAGGTTAAGACATGGTATCAAAATCGTCGAATGAAGTGGAAGAAAGAAACATTTGAAAGTGAAGAT 330
Q  V  K  T  W  Y  Q  N  R  R  M  K  W  K  K  E  T  F  E  S  E  D   110
AAAGAACCAAAGATATCGTAAAAAATTGTCAGTGAAGTGACCGTATGTAGTTTCCGGAAGCAAACT 396
K  E  P  K  I  S  *                                                 116
CTTTTGAAATATCTGCTACAAAAAAAACCAGTCTTGTGGGTCAGTTTCAAATAAAAAGAAGCACAC 462
TAGAGTTTTTATGAAAAAAAAACGATCATTAAAATAAAGACAATCAAACAATTTTCGATAGAGTTT 528
AAAAGTTTTACAAATTAGCTGTACTACGTACCGTTACTACGTACAACTATGATCTACTGATCGAAG 594
GAACTTTCAATTCTTATTAACCATAATGAATAAAGAAAACATGTTTCGCAGCAAATTAAATATTTT 660
GAATAACGGGTGATGTAAATGAGTGATGTAG                                    691
```

cnox4

```
 AATTAATGCTTACCGATAATCAGGTGAAAACATGGTATCAAAATCGTCGAATGAAGTTAAAGCGT  65
    L  M  L  T  D  N  Q  V  K  T  W  Y  Q  N  R  R  M  K  L  K  R     21
CACATAAAACTTCATATGCAACATTGTCGTGAGAATTTAAGCAACCGAAGTGTTGCTAATTATTTT 123
H  I  K  L  H  M  Q  H  C  R  E  N  L  S  N  R  S  V  A  N  Y  F     43
CGACCTCATGTACTTACTATTGCTGAAGAGAATGCGTACAAAGACTTTTCAGAAGGCCATCTTTGA 186
R  P  H  V  L  T  I  A  E  E  N  A  Y  K  D  F  S  E  G  H  L  *     64
TAAATGCAATTGCCGGAAGAGGTATTTTCCAGCAAGAAAGAACCTGAAAAATGCACAAATATACAT 252
AACTGCACATATATGTAAATACTTTATAAAAATGTTTGTAAATATAAGAAAACAGCAAAATAACTT 318
ATGAGATTATGTTTTAACAAAGAACAAGATGATGATGATGATAATGAT                   377
```

msh

```
AAGAGTTTCAATTTGATCTGAGCAAATGCTTCCTTCGTAAGCACAAAGCTAATCGCAAGCCACGG  65
  E  F  Q  F  D  L  S  K  C  F  L  R  K  H  K  A  N  R  K  P  R     21
ACACCTTTTCAGTTAATCAACTTCTCACACTCGAACAAAAGTTCAAGCGAAAACAATATTTATCA 131
T  P  F  S  V  N  Q  L  L  T  L  E  Q  K  F  K  R  K  Q  Y  L  S    43
ATATCAGAAAGAGCGGAATTAAGCGAATTACTAAGGCTTACAGAAACTCAAATAAAAATCTGGTTT 197
I  S  E  R  A  E  L  S  E  L  L  R  L  T  E  T  Q  I  K  I  W  F    65
CAAAATCGACGTGCAAAACAAAAGAGAAGCAAAGAAGCCGAGATAGAGGAGTCAGTTCGAAACCGT 263
Q  N  R  R  A  K  Q  K  R  S  K  E  A  E  I  E  E  S  V  R  N  R    87
TTGCCCCTTTCAGCAGCAGATTATCGATCTTTGGATCATCTCACGTTATTATCCAGTTTCATCGCG 329
L  P  L  S  A  A  D  Y  R  S  L  D  H  L  T  L  L  S  S  F  I  A   119
TTTATCCCAATCGAGTATGAAGTAAAAATATTCATGAATGTTCAACGAGGCATAGAATAGTTTTCG 395
F  I  P  I  E  Y  E  V  K  I  F  M  N  V  Q  R  G  I  E  *          138
TAAAAAGTATTTCATCAGGAGCTTCGTCTTCAGTTTACCTCAAGTTAATCTTTCATCAGCCACAAA 461
TTAAAACAAACCTCTATCAAACTGTTTTGTATATATCTTATATGTTAATTGTAAATAGTTACTCTT 527
GCCG                                                               531
```

Fig. 3. DNA and predicted protein sequences of *cnox1*, *cnox2*, *cnox3*, *cnox4* and *msh* genes. In each case the homeodomain and the putative polyadenylation signals have been underlined. The LYP motif in *cnox2* and *cnox3* sequences is shadowed.

N-terminal region (Figure 4). Outside these regions only weak homologies were observed (not shown), and the C-terminal part of *cnox2* is serine-rich as observed in most *Dfd* type homeoproteins (Mavilio *et al.*, 1986; Featherstone *et al.*, 1988; Galliot *et al.*, 1989). A long 3′ AT-rich untranslated sequence common for homeobox genes is present for most of the hydra genes but none of the clones contained a poly(A) tail (Figure 3).

On a genomic Southern blot *cnox1*, *cnox3* and *msh* show unique major bands (Figure 5A); *cnox2* exhibits some cross hybridization with one or two *C. viridissima*-related sequences as well as to its *Hydra vulgaris* homologue. We would therefore consider *cnox1*, *cnox2*, *cnox3* and *msh* to be single copy genes. Under our hybridization conditions, no cross hybridization could be detected with rat genomic DNA; genomic DNA extracted from *Artemia salina* (which was used as hydra food) was hybridized to the *cnox1*, *cnox2* and *msh* probe as a control and gave no signal (data not shown).

Expression of cnox genes in regenerative versus non-regenerative conditions

Northern blot analysis revealed unique transcripts for *cnox2* (1.4 kb), *cnox3* (1.9 kb), *msh* (1 kb) and *cnox1* (0.9 kb, Figure 5B and not shown). In order to measure expression levels of these genes along the body column and during head regeneration, we took advantage of the quantitative PCR technique to test the same cDNA template with five different genes (see below). This technique requires small amounts of RNA, is very sensitive and allows comparative studies. However, the changes in transcript abundance detected by this experimental approach can be explained in terms of either messenger stability modification or transcriptional regulation.

Total RNAs isolated from either heads, total animals or gastric regions at different time points during regeneration (see Figure 6) were reverse transcribed after DNase treatment. The resulting cDNAs were then subjected to the polymerase chain reaction (PCR) using four different pairs of oligonucleotides corresponding to the four hydra homeobox genes (Figure 7a). Control experiments were carried out to verify the absence of DNA contamination in the RNA and to select, for each gene, the number of cycles corresponding to the exponential phase of the curve (Figure 7b). We also ensured that the amount of PCR product was proportional to the amount of RNA template (Figure 7c).

However, because of the differences in efficiency of both the reverse transcription and the annealing of the different sets of primers, the absolute amounts of PCR products cannot be accurately compared from one gene to another. Finally we used the hydra *Actin Associated Protein* gene (*AAP*, Keppel and Schaller, 1991) as a control for the variations of expression of an unrelated gene.

The level of transcripts per microgram of total RNA in uncut animals was in the same range as that measured in heads, being in most cases slightly higher. The lower levels in cut animals might be due to transient mRNA degradation induced by cutting. For all homeobox genes tested, as well as for the non-homeobox gene, *AAP*, hydra heads contain a few more transcripts per microgram of total RNA than gastric regions when RNAs are isolated at time zero (Figure 7d–h). This distribution of transcripts along the body axis may therefore not be specific for homeobox genes and could reflect a more intense transcriptional activity for a wide variety of genes in heads. This head–gastric gradient, however, is steeper for *cnox3* (Figure 7g) and *cnox4* (data not shown).

Fig. 4. Conserved stretches of homology outside the homeodomain between the *cnox2* protein and *Dfd*-type homeoproteins: Identical or conservatively substituted amino acids are represented on a grey background. Spaces are used to optimize the alignment. The N-terminal sequence of *cnox2* is aligned with that of *Dfd*-type homeoproteins. The hexapeptide motif (VYPWMK) found in *Dfd*-type homeoproteins is aligned with the *cnox2* sequence; the large triangle represents the start of the homeodomain.

Gastric slices from the multiheaded mutant of *C. viridissima* regenerate heads on both ends but hardly ever regenerate a foot. After 24 h, tentacle buds become visible; after 50 h a complete, two-headed green hydra has formed (Figure 6d and e). Two hours after cutting, a drastic drop in transcript abundance was observed in all cases, including *AAP* but excluding *cnox3*. The lack of a drop in *cnox3* transcript level could be explained by the initial very low level of expression in the gastric column. Subsequently, clear differences in patterns of expression of the different genes were noticed. First, the level of *AAP* transcripts drastically increased between 2 and 4 h after cutting, reaching a steady plateau in the range of the total uncut value, ~50% above the initial head value (Figure 7d). This plateau was maintained for the following 2 days. The transcript abundance after 5 days was quite low, as in all genes tested, and is likely to be due to a decrease in mRNA caused by the lack of feeding of the animals during regeneration. Secondly, *cnox1* transiently showed a very high level of transcripts within the first hours of head regeneration, exceeding the normal head level by ~100% (Figure 7e). By 24 h after cutting, the *cnox1* transcript level was back to the initial head value and remained stable on the following day. Thirdly, 4 h after cutting, *cnox2* transcripts rose to a value ~50% above the initial head value but lower than the 24 h value. *cnox2* expression then slowly declined over the next days (Figure 7f). Fourthly, in contrast to *cnox1* and *cnox2*, the *cnox3* RNA level remained low during the first hours and increased only later to reach the initial head value after 24 h, doubling then this value after 2 days (Figure 7g). Finally, *msh*, after the initial drop, went back to the initial head value 4 h after cutting and remained at that level for the following 2 days (Figure 7h). A second different regeneration experiment gave similar results.

The differential expression of these five genes during head regeneration, as depicted in Figure 7i, suggest different specific requirements for several of these gene products: early and at a high level in the case of *cnox1*, early and prolonged during the first day in the case of *cnox2* and

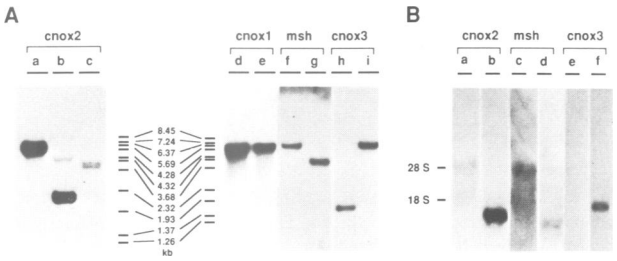

Fig. 5. (**A**) Genomic Southern blots. 10 μg of genomic DNA were digested with *Eco*RI (lanes a, c, d, f and h) or *Hin*dIII (lanes b, e, g and i) and loaded onto each lane. Lanes a, b and d–i: *C. viridissima*; lane c: *Hydra vulgaris*. (**B**) Northern blots: lanes a, c and e: 20 μg poly(A)$^-$ RNA; lanes b and d: 6 μg poly(A)$^+$ RNA; lane f: 4 μg poly(A)$^+$ RNA; all RNAs are from *C. viridissima*; in the case of *msh* the specific signal (lane d) could be obtained only by an exposure time twice as long as that for the other genes.

delayed in the case of *cnox3*. *AAP* and *msh* are expressed at high and low levels respectively during the regenerative processes but do not exhibit any significant modulation.

Discussion

Hydra has HOM/HOX ancestor genes

We used an *Antp*-derived guessmer oligonucleotide and isolated five different homeobox genes from hydra. Two of these, *cnox1* and *cnox2*, are related to the HOM/HOX genes. This indicates that HOM/HOX genes evolved from common ancestors already present in diblastic animals. It should be mentioned that none of the hydra homeodomains are significantly more conserved with either arthropod or vertebrate homeodomains.

The hydra *cnox1* sequence exhibits some characteristic features of the *lab* homeodomain family, but also displays comparable levels of similarity with the *Dfd* and *AbdB* homeodomain families. These three types of homeodomains might derive from a common ancestor to which *cnox1* could

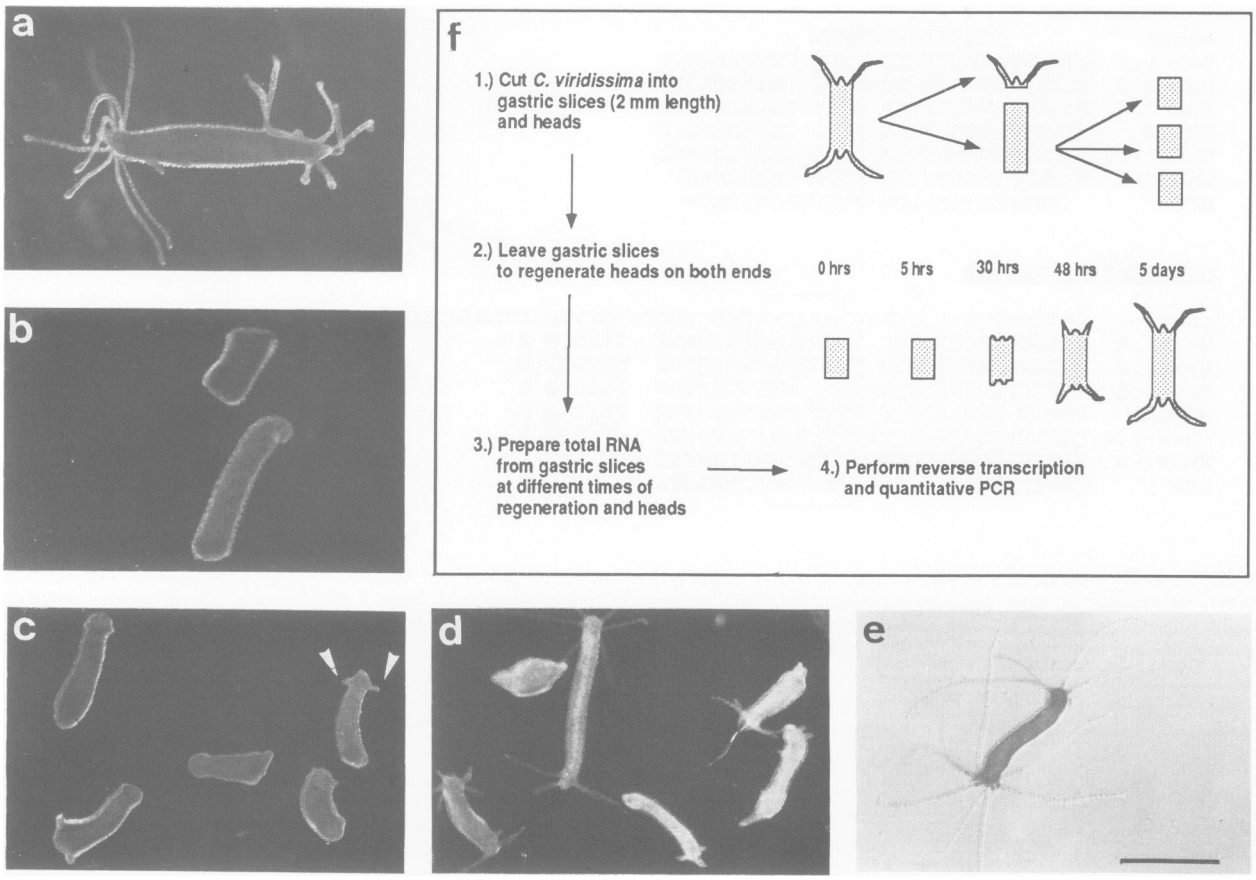

Fig. 6. Head regeneration in the multiheaded mutant of *Chlorohydra viridissima*: (a–e) Photographs of living *C. viridissima* as total uncut animal (a) and during head regeneration (b–e). The bar in (e) represents 1 mm and refers to all photographs. (**a**) Total uncut hydra with heads on both ends and an additional head budding from the gastric column. Note that for technical reasons a relatively small animal was chosen (2 mm long); the average length of such mutant hydra is 5–10 mm. (**b**) Gastric slices 30 min after cutting. (**c**) Gastric slices after 30 h of regeneration: tentacles are forming (arrowheads) at one end and, ~5 h later, also at the other end (not shown). (**d**) Gastric slices after 52 h of regeneration: most of the slices have regained an adult-like shape. (**e**) Gastric slice after 5 days of regeneration. (**f**) Scheme of the cutting procedure.

be closely related. This would be in good agreement with the predictions of Scott *et al.* (1989). The *cnox2* sequence is clearly homologous to *Dfd* type homeodomains, containing in addition common features in regions outside of the homeodomain, like in the hexapeptide and the N-terminal regions. It is worth noting that the *cnox1* and *cnox2* homeodomains are less related to each other (51% identical, 69% similar) than to their arthropod or vertebrate counterparts. Up to now vertebrate or *Drosophila* homeobox probes have failed to detect homeoboxes in lower eukaryotes (McGinnis, 1985; our work). In hydra this is probably due to different codon usage. As the *cnox1* and *cnox2* homeoboxes are >50% identical to *Antp*, closer *Antp*-related hydra homeoboxes would have been detected with the *Antp*-like probes under the low stringency conditions we used. Hence we cannot expect a higher degree of conservation at the DNA level between higher and lower eukaryotic homeoboxes, and the presence in hydra of homeodomains more related to the *lab* or *Dfd* types seems unlikely. *cnox1* and *cnox2* might therefore represent the genuine ancestors of *lab* and *Dfd* family members. Their chromosomal organization is currently under investigation.

One of the hydra homeodomains (*msh*) resembles the *msh/Hox-7* type homeodomain, particularly in variable regions (e.g. the first 10 amino acids) and in flanking regions of the homeodomain. Several *msh* gene duplications occurred in the lineage leading to vertebrates, but only one *msh* gene has been identified in ascidia and *Drosophila* (Holland, 1991). Therefore it is supposed that the hydra *msh* gene is unique. The high conservation (>70% identical) might seem surprising since *msh/Hox-7* genes show a different pattern of expression from other *Hox* genes. This pattern is consistent with a role for *msh/Hox-7* gene products in the subdivision of mesoderm lineage and in the meso-ectodermal inductive interactions that result in pattern formation (Hill *et al.*, 1989; Robert *et al.*, 1989; Bodmer *et al.*, 1990; Su *et al.*, 1991; Suzuki *et al.*, 1991; Yokouchi *et al.*, 1991). As mesoderm does not exist in diblastic animals, this could indicate that the particular involvement of *msh/Hox-7* in mesoderm is the result of a late recruitment from another, older function.

The last hydra homeodomain classified is *cnox3* which is most similar to the *BarH1* and *Dll* type homeodomains. In *cnox3* and *cnox4* the invariant Phe50 residue is substituted by a Tyr. In *engrailed*, this residue is important in stabilizing the overall fold of the homeodomain by interacting with the Glu21 of helix 1 (Wolberger *et al.*, 1991). As Glu21 is present in *cnox3* and as tyrosine is not significantly bigger than phenylalanine, this interaction could be weakened but not drastically modified.

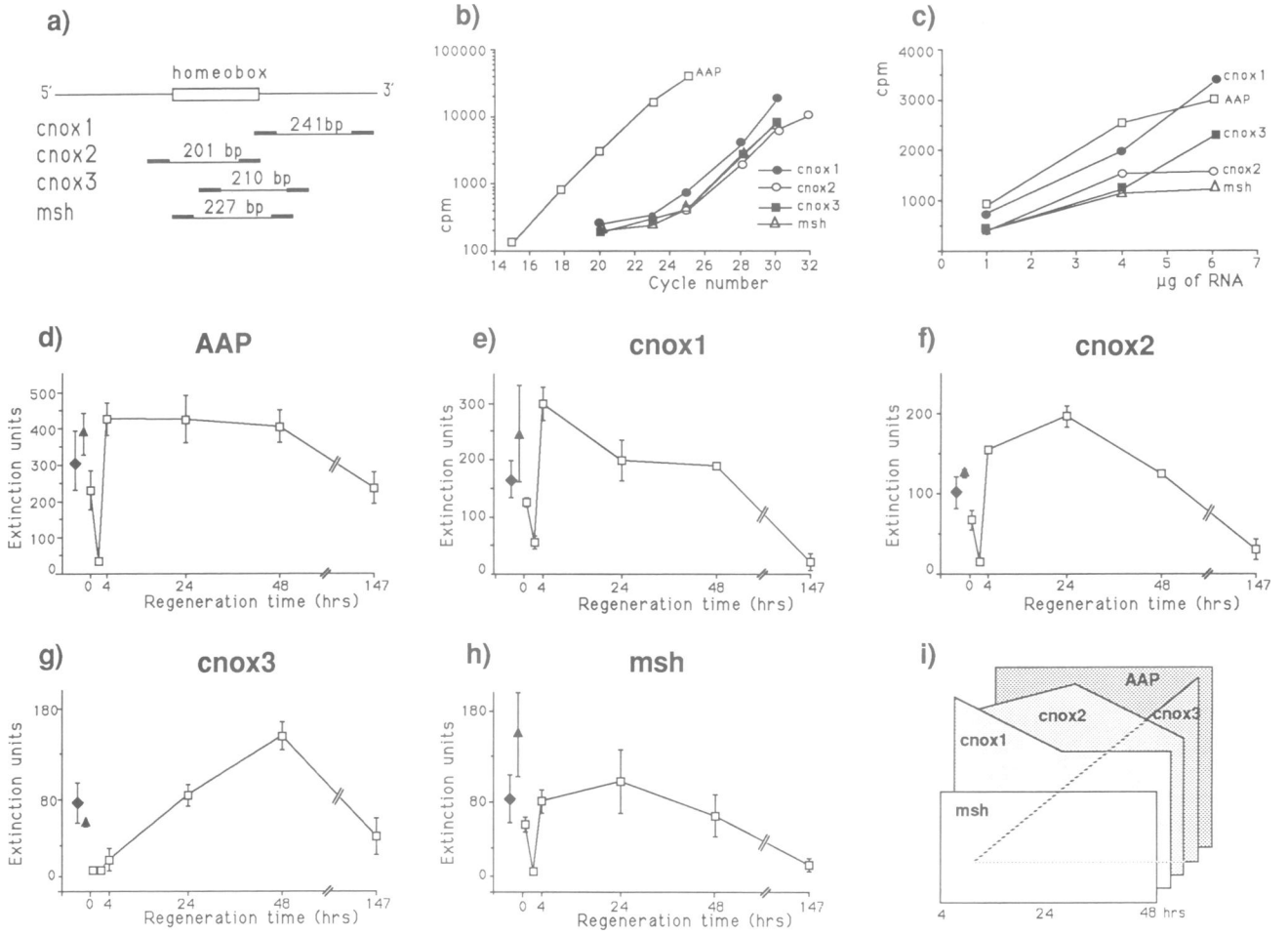

Fig. 7. Quantitative PCR of hydra homeobox genes. (**a**) Pairs of primers used for each gene. The four oligonucleotide pairs and their positions relative to the homeodomain are indicated by black bars, the numbers give the length of the final PCR product. (b–h) RNA levels determined as incorporation of [α-^{32}P]dCTP measured in PCR products (c.p.m. or extinction values). (**b**) Kinetics of quantitative PCR. Increasing numbers of cycles were performed with 1 μg of reverse transcribed total RNA. For each gene a cycle number in the logarithmic phase was chosen for the subsequent experiments: 18 cycles for *AAP* and 26 cycles for *msh*, *cnox1*, *cnox2* and *cnox3*. (**c**) PCR products are proportional to the amount of total RNA used in reverse transcription reaction (1, 4 and 6 μg). (**d–h**) Variations in transcript amount along the body axis and during head regeneration. The name of each investigated gene is indicated above the corresponding graph. Black triangles correspond to total uncut animals, black diamonds to heads and unfilled squares to gastric slices. (**i**) Scheme summarizing the expression pattern of the five tested genes. Although the absolute values cannot be strictly compared from one gene to another, the variations of these values along the time axis exhibit some clear cut differences between these genes, especially at the time points of 4, 24 and 48 h which are shown here.

Differential expression of cnox genes during regeneration

A feature common to all genes investigated is that there are slightly more transcripts in uncut animals than in cut heads or cut gastric columns. Time-point zero RNAs were prepared from heads or gastric regions within 30 min following cutting. Thus comparisons between expression levels in either cut or uncut animals are limited by our lack of knowledge about the early molecular consequences of the cutting injury. Lower transcript levels in heads and gastric regions at time zero could be only due to an immediate start of the drastic drop in messenger abundance which we observed 2 h after cutting. Additionally, in all genes we have examined, we found that the initial head value was higher than the initial gastric one, this gradient seem to vary from one gene to the other.

The significant drop in gastric transcript abundance, detected 2 h after cutting, may be correlated with the inhibition of mitosis and DNA replication observed during the first hours of regeneration (Park *et al.*, 1970). This might reflect a short and general decrease in metabolic activity. A strong increase in sequence-specific DNA binding activity onto the cAMP response element is observed 2 h after cutting, which may indicate that some transcriptional processes start to be enhanced when mRNA levels are still very low (B.Galliot, unpublished results).

After the initial drop, the *AAP* expression level remained on a plateau above the initial head value, indicating that this gene's product is required through the whole regenerative process. The four homeobox genes are differentially expressed during head regeneration of gastric slices (Figure 7d–i). Four hours after cutting, *cnox1* displays a peak which is ~2-fold more than the initial head value. It may be correlated with the major cellular events occurring at this time, i.e. the head determination and an increase in nerve cell differentiation (Hoffmeister and Schaller, 1987). An elevated amount of *cnox2* transcripts is also observed at this time, although to a lesser extent, a peak being reached 24 h after cutting. This pattern implies that the *cnox2* gene product is particularly involved at this period and at twice

the initial head value. The *cnox3* expression increases from a low level immediately after cutting to twice the initial head value after two days, suggesting that the *cnox3* gene product may be required either for late events only (e.g. head-specific nerve or epidermal cell differentiation) or for the maintenance of head differentiation. Finally *msh* exhibits weak modulations at a low level of expression, indicating that it plays no essential role during head regeneration. Taking the initial head value as a reference, *cnox1*, *cnox2* and *cnox3* display their maximal levels of expression at different stages of head regeneration, as summarized in Figure 7i.

If we consider that *cnox1* and *cnox2* correspond to the hydra *lab/Hox-1.6* and *Dfd/Hox-1.4* cognates respectively, and that early strong expression during head regeneration correlates with a specific function for head differentiation, it then appears that according to their expression patterns these HOM/HOX-related genes are differentially required along the anteroposterior axis. Thus they would follow the rules observed for the cognate genes in arthropods and vertebrates, i.e. that *lab/Hox-1.6* family members are involved more anteriorly than *Dfd/Hox-1.4*-like genes (Lufkin *et al*, 1991).

In hydra, head and foot differentiation are regulated by a set of substances acting as extracellular signals (for review see Schaller *et al.*, 1989). Since *cnox* genes display different patterns of expression during head regeneration, it will be of interest to know which of these patterns are dependent on these extracellular signals and/or are involved in their regulation.

Materials and methods

cDNA library, cloning strategy and sequencing

Homeobox genes were isolated from an oligo(dT)-primed λgt10 *C. viridissima* cDNA library built using a Pharmacia kit: double-stranded cDNA was synthesized from 4 μg of poly(A)$^+$ RNA, after ligation to *Eco*RI digested λgt10 arms (10 ng of cDNA for 330 ng of vector) the phage DNA was packaged (Stratagene) resulting in 1.5×10^7 primary plaques per microgram of cDNA. The library was amplified up to 5×10^{10} recombinants. After transfer to nitrocellulose filters, 5×10^5 clones were screened with an end-labelled 50-guessmer oligonucleotide (5×10^5 c.p.m./ml) under low stringency conditions (hybridization: 6% formamide, 6×SSC, 5×Denhardt's, 100 μg/ml tRNA, 50 mM sodium pyrophosphate and 1% SDS; washes: 6×SSC and 0.1% SDS twice for 10 min at room temperature, 2×SSC and 0.1% SDS twice for 10 min at room temperature). The 50mer sequence (ACT GAA A/CGI CAA ATA AAA ATA TGG TTT CAA AAT A/CGI A/CGI ATG AAA TGG AA) corresponds to the sequence of the third helix of the *Antp* homeodomain (amino acids 42–56, Figure 1). Three inosine residues were included to decrease the redundancy further to 8-fold. Seven positive plaques were purified from a second round of screening, cDNA inserts were subsequently subcloned into the pBS vector (Stratagene) and sequenced by the dideoxy chain termination method (Sanger *et al.*, 1977). Computer analysis of sequences was performed using the Heidelberg University Sequence Analysis Resource (HUSAR).

Southern and Northern analyses

Genomic DNA from 2000 animals (*C. viridissima, Hydra vulgaris* or *Hydra oligactis*) was prepared according to Stafford (1976), 20 μg of digested DNA were separated on a 1% agarose gel, denatured and transferred to Hybond membranes (Amersham procedure). Gel-purified cDNA inserts were labelled by random priming (Amersham kit) and hybridized for 2 days in 35% formamide, 5×SSC, 0.1% SDS, 5×Denhardt's, 500 μg/ml salmon sperm DNA and 100 μg/ml tRNA at 42°C and washed for 2×10 min in 5×SSC, 0.1% SDS, then for 10 min in 2×SSC and 0.1% SDS, both at room temperature.

Total RNA from 2000 animals that had been starved for 2 days was prepared according to Chomcyzynski and Nicoletta (1987) and poly(A)$^+$ RNA was separated using oligo(dT)–cellulose type 7 (Pharmacia). 20 μg of total RNA and 6 μg of poly(A)$^+$ RNA were separated on a formaldehyde gel (Sambrook *et al.*, 1989) and, after transfer onto Hybond membrane, were hybridized to labelled cDNA inserts for 2 days in 50% formamide, 5×SSPE, 1% SDS, 5×Denhardt's, 100 μg/ml tRNA at 42°C. Washes were carried out for 2×10 min in 2×SSC, 1% SDS, then for 10 min in 0.2× SSC, 1% SDS, both at room temperature.

Culture of animals and regeneration procedure

C. viridissima was cultured in hydra medium consisting of 1 mM $CaCl_2$, 0.1 mM KCl, 0.1 mM $MgCl_2$ and 0.5 mM NaH_2PO_4/Na_2HPO_4 pH 7.6. The animals were fed daily between 09.00 and 10.00 h with nauplii of *Artemia salina* and washed 6 h later. Before cutting, animals were starved for 24 h. 300 animals for each time point were then cut just below the hypostome, pure gastric regions were transferred to Petri dishes containing 30 ml of hydra medium and divided into slices of 1–2 mm length. The medium was changed every 2 days. Time point zero corresponds to tissues, heads and gastric columns collected within the first 30 min after cutting.

Quantitative PCR

RNAs from each time point were prepared according to Gough *et al.* (1988). 10 μg of total RNA were digested with 5 U RNase-free DNase I (Boehringer) for 15 min at 37°C. The enzyme was inactivated by phenol–chloroform extraction and, after ethanol precipitation, the amount of total RNA was calculated from its OD_{260} value. Reverse transcription was performed on 1, 4 and 6 μg of total RNA hybridized to an oligo(dT) oligonucleotide (17mer), 0.5% of the resulting cDNA was used as template for the PCR reaction performed in 25 μl of 10 mM Tris–HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$ and 0.01% gelatin in the presence of 0.5 μM of each primer and 1 U *Taq* polymerase (Amplitaq, Perkin–Elmer Cetus) with the following profile: 1 min 94°C, 40 s 94°C/40 s 55°C/40 s 72°C for four cycles, 30 s 90°C/30 s 55°C/40 s 72°C for 18–26 cycles. By using non-reverse transcribed total RNAs as template, no PCR products were detected (data not shown). The T_m of each oligonucleotide was calculated according to the method of Lathe (1985) to be equal to 70°C. The sequences of the oligonucleotides used are as follows: cnox1–5′: cttagtaaatgaccgccgggttaa; cnox1–3′: aaacgactcattacgattgggtct; cnox2–5′: atgggggaatagctatatctttctta; cnox2–3′: tcaaaacggatccgaaccggcta; cnox3–5′: ttgaaactgacccacaagactggt; cnox3–3′: tagcagaccgacttggcctgaat; msh1–5′: aggggcaaacggtttcgaactga; msh1–3′: caaagctaatcgcaagccacgga; AAP-5′:ccagtcagcgcctgcagcaga; AAP-3′: tgtaacacggattcgcctttgaga. (The oligonucleotides were synthesized in R.Frank's laboratory, ZMBH.)

In each reaction 1 μCi of [α-^{32}P]dCTP (Amersham, 3000 mCi/mmol) was included, the samples were then separated on a 6% polyacrylamide gel, which was subsequently dried and exposed to autoradiography. The respective bands were cut out and measured by scintillation counting (Figure 7b and c) or the X-ray film was scanned on an Elscript 400 scanning machine (Hirschmann); the area of a peak corresponding to a PCR band was calculated and given as extinction units (Figure 7d–h). As the amount of PCR product is proportional to the amount of template only during the logarithmic phase of a PCR (Chelly *et al.*, 1990; Robinson and Simon, 1991), five PCRs were performed with increasing numbers of cycles for each pair of primers onto one cDNA (from the head fraction). Thus, for each gene a specific number of cycles for a specific amount of RNA template was determined and used for the following experiments. Three identical PCRs were performed and the average values and standard deviations were plotted on Figure 7.

Acknowledgements

We are very grateful to Sylvia Kreger, Birgitta Schinke and Stefanie Schroeder for their daily care of the hydra culture and their help in the cutting procedure. We also want to thank Denis Duboule for his critical reading of the manuscript. This work was supported by grants from the Deutsche Forschungsgemeinschaft (SFB 317), the Bundesministerium für Forschung und Technologie (BCT 365/1) and the Fonds der Deutschen Chemischen Industrie.

References

Akam,M. (1989) *Cell*, **57**, 347–349.
Beeman,R.W. (1987) *Nature*, **327**, 247–249.
Bodmer,R., Jan,L.Y. and Jan,Y.N. (1990) *Development*, **110**, 661–669.
Bogarad,L.D., Utset,M.F., Awgulewitsch,A., Miki,T., Hart,C.P. and Ruddle,F.H. (1989) *Dev. Biol.*, **133**, 537–549.
Boncinelli,E., Simeone,A., Acampora,D. and Mavilio,F. (1991) *Trends Genet.*, **7**, 329–334.
Bosch,T.G., Unger,T.F., Fisher,D.A. and Steele,R.E. (1989) *Mol. Cell. Biol.*, **9**, 4141–4151.

Bürglin,T.R., Finney,M., Coulson,A. and Ruvkun,G. (1989) *Nature*, **341**, 239–243.
Bürglin,T.R., Ruvkun,G., Coulson,A., Hawkins,N.C., McGhee,J.D., Schaller,D., Wittmann,C., Muller,F. and Waterston,R.H. (1991) *Nature*, **351**, 703.
Chelly,J., Montarras,D., Pinset,C., Berwald-Netter,Y., Kaplan,J.C. and Kahn,A. (1990) *Eur. J. Biochem.*, **187**, 691–698.
Chomcyzynski,P. and Nicoletta,S. (1987) *Anal. Biochem.*, **162**, 156–159.
Cohen,S.M., Brönner,G., Küttner,F., Jürgens,G. and Jäckle,H. (1989) *Nature*, **338**, 432–434.
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.
Dollé,P. and Duboule,D. (1989) *EMBO J.*, **8**, 1507–1515.
Duboule,D. and Dollé,P. (1989) *EMBO J.*, **8**,1498–1505.
Fainsod,A. and Gruenbaum,Y. (1989) *FEBS Lett.*, **250**, 381–385.
Featherstone,M.S., Baron,A., Gaunt,S.J., Mattei,M.G. and Duboule,D. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 4760–4764.
Fisher,D.A. and Bode,H.R. (1989) *Gene*, **84**, 55–64.
Frohman,M.A., Boyle,M. and Martin,G.R. (1990) *Development*, **110**, 589–607.
Galliot,B., Dolle,P., Vigneron,M., Featherstone,M.S., Baron,A. and Duboule,D. (1989) *Development*, **107**, 343–359.
Garcia-Fernandez,J., Baguna,J. and Salo,E. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 7338–7342.
Gehring,W.J. (1987) In Firtel,R.A. and Davidson,E.H. (eds), *Molecular Approaches to Developmental Biology*. Liss, New York, pp. 115–129,
Gehring,W.J., Müller,M., Affolter,M., Percival-Smith,A., Billeter,M., Qian,Y.Q., Otting,G. and Wüthrich,K. (1990) *Trends Genet.*, **6**, 323–329.
Gough,N.M. (1988) *Anal. Biochem.*, **173**, 93–95.
Graham,A., Papalopulu,N., and Krumlauf,R. (1989) *Cell*, **57**, 367–378.
Hill,R.E., Jones,P.F., Rees,A.R., Sime,C.M., Justice,M.J., Copeland,N.G., Jenkins,N.A., Graham,E. and Davidson,D.R. (1989) *Genes Dev.*, **3**, 26–37.
Hoffmeister,S.A.H. and Schaller,H.C. (1987) *Dev. Biol.*, **22**, 72–77.
Holland,P.W.H. (1991) *Gene*, **98**, 253–257.
Kaufman,T.C., Lewis,R. and Wakimoto,B.T. (1980) *Genetics*, **94**, 115–133.
Keppel,E. and Schaller,H.C. (1991) *J. Cell Sci.*, **100**, 789–797.
Kim,Y. and Nirenberg,M. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 7716–7720.
Kojima,T., Ishimaru,S., Higashijima,S., Takayama,E., Akimaru,H., Sone,M., Emori,Y. and Saigo,K. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 4343–4347.
Lathe,R. (1985) *J. Mol. Biol.*, **183**, 1–12.
Lewis,E.B. (1978) *Nature*, **276**, 565–570.
Lufkin,T., Dierich,A., LeMeur,M., Mark,M. and Chambon, P. (1991) *Cell*, **66**, 1105–1119.
Mavilio,F. *et al.* (1986) *Nature*, **24**, 664–667.
McGinnis,W. (1985) *Cold Spring Harbor Symp. Quant. Biol.*, **50**, 263–270.
Monaghan,A.P., Davidson,D.R., Sime,C., Graham,E., Baldock,R., Bhattacharya,S.S. and Hill,R.E. (1991) *Development*, **112**, 1053–1061.
Murtha,M.T., Leckman,J.F. and Ruddle,F.H. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 10711–10715.
Park,H.D., Ortmeyer,A.B. and Blankenbaker,D.P. (1970) *Nature*, **227**, 617–619.
Porteus,M.H., Bulfone,A., Ciaranello,R.D. and Rubenstein,J.L.R. (1991) *Neuron*, **7**, 221–229.
Price,M., Lemaistre,M., Pischetola,M., Di Lauro,R. and Duboule,D. (1991) *Nature*, **351**, 748–751.
Rangini,Z., Frumkin,A., Shani,G., Guttmann,M., Eyal-Giladi,H., Gruenbaum,Y. and Fainsod,A. (1989) *Gene*, **76**, 61–74.
Robert,B., Sassoon,D., Jacq,B., Gehring,W. and Buckingham,M. (1989) *EMBO J.*, **8**, 91–100.
Robinson,M. and Simon,M.I. (1991) *Nucleic Acids Res.*, **19**, 1557–1562.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sanger,F.S., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Sasaki,H., Yokoyama,E. and Kuroiwa,A. (1990) *Nucleic Acids Res.*, **8**, 1739–1747.
Schaller,H.C., Hoffmeister,S.A.H. and Dübel,S. (1989) *Development*, **107** (suppl.), 99–107.
Schaller,D., Wittmann,C., Spicher,A., Müller,F. and Tobler,H. (1990) *Nucleic Acids Res.*, **18**, 2033–2036.
Schierwater,B., Murtha,M., Dick,M., Ruddle,F.H. and Buss,L.W. (1991) *J. Exp. Zool.*, **260**, 413–416.
Scott,M.P., Tamkun,J.W. and Hartzell,G.W. (1989) *Biochim. Biophys. Acta Rev. Cancer*, **989**, 25–48.
Stafford,B. (1976) *Nucleic Acids Res.*, **3**, 2303.
Su,M.W., Suzuki,H.R., Solursh,M. and Ramirez,F. (1991) *Development*, **111**, 1179–1187.
Sundin,O.H., Busse,H.G., Rogers,M.B., Gudas,L.J. and Eichele,G. (1990) *Development*, **108**, 47–58.
Suzuki,H.R., Padanilam,B.J., Vitale,E., Ramirez,F. and Solursch,M. (1991) *Dev. Biol.*, **148**, 375–388.
Takahashi,Y. and Le Douarin,N. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 7482–7486.
Wang,G.V.L., Dolecki,G.J., Carlos,R. and Humphreys,T.(1990) *Dev. Genet.*, **11**, 77–87.
Wolberger,C., Vershon,A.K., Liu,B., Johnson,A.D. and Pabo,C.O. (1991) *Cell*, **67**, 517–528.
Yokouchi,Y., Ohsugi,K., Sasaki,H. and Kuroiwa,A. (1991) *Development*, **113**, 431–444.

Received on December 23, 1991; revised on February 10, 1992

Note added in proof

The sequence data reported here have been deposited in the EMBL Data Library under the accession numbers: X64625, cnox1; X64626, cnox2; X64627, cnox3; X64628, cnox4; and X64629, msh.