Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 7227–7231, August 1995
Evolution

Identification of planarian homeobox sequences indicates the antiquity of most Hox/homeotic gene subclasses

(platyhelminth/Antennapedia-like genes/zootype/phylogeny)

GUILLAUME BALAVOINE AND MAXIMILIAN J. TELFORD*

Laboratoire de Biologie Cellulaire 4, Unité de Recherche Associée 1134 du Centre National de la Recherche Scientifique, Bâtiment 444, Université Paris-sud, 91405 Orsay Cedex, France

Communicated by James W. Valentine, University of California, Berkeley, CA, March 14, 1995

ABSTRACT **The homeotic gene complex (HOM-C) is a cluster of genes involved in the anteroposterior axial patterning of animal embryos. It is composed of homeobox genes belonging to the Hox/HOM superclass. Originally discovered in *Drosophila*, Hox/HOM genes have been identified in organisms as distantly related as arthropods, vertebrates, nematodes, and cnidarians. Data obtained in parallel from the organization of the complex, the domains of gene expression during embryogenesis, and phylogenetic relationships allow the subdivision of the Hox/HOM superclass into five classes (*lab, pb/Hox3, Dfd, Antp*, and *Abd-B*) that appeared early during metazoan evolution. We describe a search for homologues of these genes in platyhelminths, triploblast metazoans emerging as an outgroup to the great coelomate ensemble. A degenerate PCR screening for Hox/HOM homeoboxes in three species of triclad planarians has revealed 10 types of Antennapedia-like genes. The homeobox-containing sequences of these PCR fragments allowed the amplification of the homeobox-coding exons for five of these genes in the species *Polycelis nigra*. A phylogenetic analysis shows that two genes are clear orthologues of *Drosophila* labial, four others are members of a *Dfd/Antp* superclass, and a seventh gene, although more difficult to classify with certainty, may be related to the *pb/Hox3* class. Together with previously identified Hox/HOM genes in other flatworms, our analyses demonstrate the existence of an elaborate family of Hox/HOM genes in the ancestor of all triploblast animals.**

A subfamily of the homeodomain transcriptional regulators, defined as the Hox/HOM superclass, groups all the genes of the *Drosophila* and vertebrate homeotic gene complexes (HOM-Cs) (for review, see ref. 1) and is involved in anteroposterior axial patterning. Hox/HOM genes have been found in all the major "higher" metazoan phyla analyzed to date (2–5) but also more recently in "lower" metazoan phyla, such as nematodes [*Caenorhabditis elegans* possesses a cluster of four genes (6)], platyhelminths (7–9), and cnidarians (10–12). Their function, as has been inferred from data including classical *Drosophila* genetics and molecular techniques in flies and mice, is to determine the identity of a body section of the embryo, be it an arthropod metamere (or a group of these) or a portion of the vertebrate body axis (for review, see refs. 2, 13, and 14).

From phylogenetic analyses of the sequence data of *Drosophila* and mouse *Hox* genes, Schubert *et al.* (15) proposed that an ancestral complex of at least five Hox/HOM genes existed in the common ancestor of the coelomates (*lab, pb, Dfd, Antp*, and *Abd-B*). Furthermore, these workers proposed a complex composed of three genes at the base of the Eumetazoa, one giving rise to the 3′ genes (anteriorly expressed *lab* and *pb* cognates), another to central or medial class genes (*Dfd, Scr, Antp, Ubx*, and *abd-A* cognates), and the last to the 5′ genes (*Abd-B* cognates). They corroborated this hypothesis with the available data concerning the "lower" Metazoa (essentially *C. elegans* and the cnidarians). These data are, however, still partial, and this hypothesis will doubtless evolve with the accumulating results from various other invertebrate phyla. Bürglin (1) proposed another ancestral scenario, including the *pb*-like genes in a large *Antp* class. Taking into account these conflicting hypotheses, we use the generally accepted classification and terminology of Hox/HOM genes as being members of the *lab, pb, Dfd, Antp*, and *Abd-B* classes.

Two facts make the planarians (phylum Platyhelminthes) appropriate to study the evolution of Hox/HOM genes. Their phylogenetic position as a close but clear outgroup of the coelomates (16–19) allows inference of the state of the HOM-C complex in the ancestor of Triploblastea or Bilateria (all the Metazoa that display a bilateral symmetry, Fig. 1). In addition, some groups of planarians, particularly those belonging to the order Tricladida, display a high regeneration potential. The possibility of analyzing the expression of these genes in both normally developing and regenerating planarians gives us a tool for comparing the roles of the Hox/HOM genes in these two seemingly disparate modes of development.

To carry out as exhaustive a search as possible for Hox/HOM genes in triclad planarians, we have used the very sensitive degenerate oligonucleotide PCR method. Previous degenerate PCR screens revealed distinct *Antp*-like genes from *Phagocata*, another planarian (8). The small size of the fragments obtained prevented these authors from convincingly classifying the genes. For this reason, we used inverse-PCR to give entire homeobox sequences that allowed us to identify with far more certainty the classes represented.†

MATERIALS AND METHODS

Extraction of Genomic DNA. Whole planarians, starved for 2 weeks, were homogenized in a lysis buffer containing 100 mM Tris·HCl, pH 7.4/100 mM EDTA/0.1% SDS and incubated with proteinase K at 100 μg/ml for 5 hr at 50°C, extracted successively with phenol, phenol/chloroform, and chloroform, ethanol-precipitated in the presence of 0.3 M NaAc, and treated after resuspension with RNaseA (100 μg/ml).

Extraction of Total RNA and Reverse Transcription. Total RNA for reverse transcription-PCR was extracted from intact and regenerating planarians (2, 4, and 9 days of regeneration), using a protocol derived from Chomczynski and Sacchi (20). Three planarians for each stage of regeneration were ground in 200 μl of extraction buffer (4 M guanidinium isothiocya-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: PG*X*, paralogy group *X*; HOM-C, homeotic gene complex.

*Present address: Molecular Biology Unit, Natural History Museum, Cromwell Road, London SW7 5BD, United Kingdom.

†The sequences reported in this paper have been deposited in the GenBank data base (accession nos. L41845–L41860).

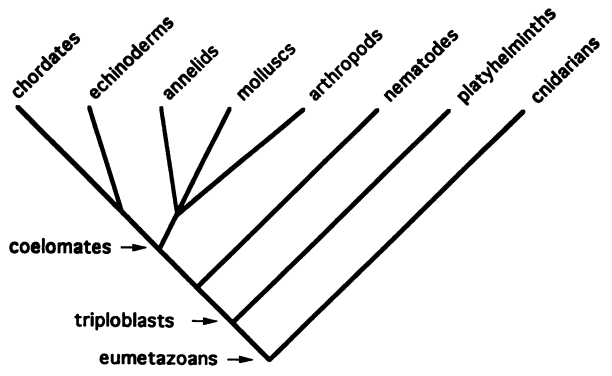

FIG. 1. Simplified phylogenetic tree of the main groups of metazoans supported by most morphological, embryological, and molecular data. This scheme corresponds, in general, to the views of Hyman (16). Coelomates are considered to be split in two lineages: protostomes (including annelids, arthropods, and molluscs) and deuterostomes (including echinoderms and chordates). The position of nematodes is not reliably determined by molecular analyses, and the point of emergence indicated, therefore, corresponds to a traditional view. The position of platyhelminths has been confirmed by molecular data (17–19).

nate/25 mM sodium citrate, pH 7.2/100 mM 2-mercaptoethanol/0.5% Sarkosyl) and agitated vigorously to shear the DNA, before addition of 25 μl of 1 M NaAc, pH 4. The sample was then extracted once with a mixture of 40 μl of phenol/10 μl of chloroform and twice with 200 μl of chloroform. Total RNA was then precipitated by addition of 3 vol of ethanol and resuspended in diethyl pyrocarbonate-treated water. Reverse transcription was done with ≈1 μg of total RNA in 25 μl of a buffer containing 50 mM Tris·HCl (pH 8.15) at 41°C, 6 mM $MgCl_2$, 40 mM KCl, 1 mM dithiothreitol, 1 mM each dNTP, 20 units of RNasin (Boehringer Mannheim), 2.5 pmol of 17-nt oligo(dT) primer, and 50 units of avian myeloblastosis virus reverse transcriptase (Boehringer Mannheim). After 2 hr of incubation at 37°C, the reaction mixture was diluted in 200 μl of 10 mM Tris, pH 7.5/1 mM EDTA.

Degenerate and Inverse PCR. For each reaction, either 50–500 ng of planarian genomic DNA or an aliquot of cDNA corresponding to 100 ng of reverse-transcribed RNA was used. The PCR primers were essentially similar to those described by Murtha *et al.* (21): The 5′ primer codes for the peptide ELEKEF present in nearly all Hox/HOM homeodomains and two alternative 3′ primers were used, one corresponding to peptide KIWFQN, present in almost all the known Hox/HOM homeodomains, and the other corresponding to the peptide KVWFQN, present in the proboscipedia-like homeodomains. Sequences for these peptides are as follows: ELEKEF, 5′-GCTCTAGARYTNGARAARGARTT-3′; KIWFQN, 5′-GGAATTCRTTYTGRAACCADATYTT-3′; and KVWFQN, 5′-GGAATTCRTTYTGRAACCANACYTT-3′. PCR cycles were done with the Bioprobe Systems (Montreuil,

D. melanogaster	*Antp*	RKRGRQTYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKEN
Presumed *lab*-class genes		
P. nigra	*Pnox2*	---Q----A------N-MG-S-T-V
P. felina	*Pfox2*	---Q----A------NSMG-S-T-V
D. lacteum	*Dlox2*	---Q----A------NSMA-S-T-V
P. woodworthii	*PWoxA*	---Q----A------TSMG-S-T-V--
P. nigra	*Pnox3*	NIT--TNF-NK-LT---------Q----A------KSMT-S-T----------Q-RRQ
P. felina	*Pfox3*	---Q----A------KSMT-S-T-V
F. hepatica	*FhHbx1*	---K----A--V---AS-Q-N-T-V--
E. trivolvis	*EToxA*	--------A------ND-G---T-V--
Presumed *Dfd*-class genes		
P. nigra	*Pnox8*	YY-Q----------KSVS-S----
D. tigrina	*Dutarh4*	Y--C--A-S-Q---------YY-Q----------NSV--S----------Y--DV
Presumed *Antp*-class genes		
P. nigra	*Pnox1a*	KR------S-H---------Q--H----------N---S----
P. nigra	*Pnox1b*	KR------S-H---------Q--H----------N---S----
P. felina	*Pfox1*	Q--H----------N---S----
D. lacteum	*Dlox1*	Q--H----------N---S----
D. tigrina	*DtHbx1*	Q--H----------N---S----
P. woodworthii	*PWoxC*	Q--H----------N---S------
P. woodworthii	*PWoxD*	Q-S----------SSN----------
E. trivolvis	*EToxC*	QYSH----------S----------
P. nigra	*Pnox7*	---K----------S-I------
D. tigrina	*Dutarh5*	---K----------T-I--------
D. tigrina	*Dutarh3*	H--S---------C---------K----------G-S----------N----DH
S. mansoni	*Smox1*	Q--T-----------------K----------T-T----------
Unclassified genes		
P. nigra	*Pnox4*	S--C-SA--NN-LV---------N--A-G--A-LSKQ-L-----V----------Q---K
D. tigrina	*Dutarh1*	TSKA-SA--NN-LV-------Y-N--A-G--T-LSKQ-L-----V----------Q---K
D. lacteum	*Dlox4*	---N--A-G--A-LSKQ-L------
P. woodworthii	*PWoxE*	---N--A-G--S-LSKQ-L--------
P. nigra	*Pnox5*	-K----S-S--L-V-KM-M-N-T-V
P. nigra	*Pnox6*	KM-L-TSFHSH-LNI-----LV-M----M----M-S--E-G-K----------VRL--IC
Non clustered *Antp*-like genes		
P. nigra	*PnMox*	AVHN----L--Y-L-V--S-N----
D. tigrina	*Dutarh6*	QRKE-TAFSKG-I-------AVHN----L--Y-L-V--N-N-----V--------C-RSK

FIG. 2. Alignments of the putative amino acid sequences of the homeobox fragments from the three planarian species considered here and related sequences from several platyhelminths obtained by other teams to date. *Drosophila* Antennapedia has been chosen as a reference sequence throughout. The table is divided into groups according to the presumed relations of homology. The unclassified sequences are those that show very derived sequences whose position in a general tree has not allowed them to be clustered unambiguously with any known group of Hox/HOM genes. The planarian sequences described here are those corresponding to the underlined species. Other sequences are from *Drosophila* (1), *Phagocata woodworthii* (*P. woodworthii*; ref. 8), *Echinostoma trivolvis* (*E. trivolvis*; ref. 8), *Dugesia tigrina* (*D. tigrina*; refs. 9, 26), *Fasciola hepatica* (*F. hepatica*; ref. 9), and *Schistosoma mansoni* (*S. mansoni*; ref. 7).

France) *Taq* polymerase as follows: 5 min at 94°C, three times (1 min at 94°C; 1 min at 60°C or 40°C; 1 min at 72°C), 27 times (1 min at 94°C, 1 min at 40°C, 1 min at 72°C), and 5 min at 72°C. DNA bands corresponding to the expected size (124 bp) were excised from a NuSieve gel (FMC) and electroeluted before cloning in pBS SK(+) vector (Stratagene). Inverse-PCR was done according to Averof and Akam (3).

Sequence Analyses. The homeodomain amino acid sequences were aligned by using the software package MUST (22). A phylogenetic tree, including the full-length platyhelminth homeodomains known to date, was constructed using the parsimony algorithm of the software PAUP 3.0 (23). PAUP distributes the sequences on a tree that links them through the shortest possible path in terms of number of mutational events required to account for the present day sequences (24). In confirmation, a very similar tree (data not shown) was calculated using NJ, the neighbor-joining algorithm (25) included in MUST, which takes into account the global similarity of the sequences. Although such tree-building methods probably fail to indicate reliable relationships between the largest subdivisions of the homeoboxes (the major subclasses mentioned in the introduction), the level of resolution is certainly sufficient to determine relationships within these major branches.

RESULTS AND DISCUSSION

Amplification of Hox/HOM Gene Fragments. PCR was done on planarian genomic DNA using both stringent and nonstringent annealing conditions. The species *Polycelis nigra* was studied extensively, whereas the two others, *Dendrocoelum lacteum* and *Polycelis felina*, were used for confirmation. The aim was to identify orthologous genes that should be very similar in three closely related species, so eliminating possible DNA contamination from unrelated species. We cloned a product of the expected size for each species, as well as larger products when using nonstringent cycling. None of the larger PCR fragments has revealed any homeobox sequence. A total of 66 clones from *P. nigra* were sequenced, sufficient to represent the diversity of amplified genes. Seven different PCR fragments were characterized (Pnox1a/b to -6). Very similar fragments were obtained from the two other species. Only the homeobox fragments Pnox5 and Pnox6 were obtained from the *P. nigra* DNA without any counterpart in *P. felina* and *D. lacteum*. Twenty-eight clones obtained from the cDNA of regenerating *P. nigra* were sequenced, adding three specific fragments (Pnox7, -8, and PnMox) together with Pnox1 and Pnox3. The flanking regions of four of the PCR fragments from *P. nigra* have been obtained either by inverse PCR or, for Pnox6, by sequencing a DNA fragment from a genomic library. No flanking regions have yet been sequenced for Pnox2 and Pnox5. Genes very similar to the cDNA fragments Pnox7, Pnox8, and PnMox have already been identified in *Dugesia tigrina* (26). We therefore focused our work on the clones specific to genomic amplifications.

The putative amino acid sequences of the different planarian homeoboxes are summarized in Fig. 2 alongside the most similar fragments obtained by other teams from other platyhelminth species. *Phagocata woodworthii* (8) and *Dugesia tigrina* (9, 26) belong to the same order, Tricladida, as *P. nigra* and *D. lacteum*. *Fasciola hepatica* and *Echinostoma* are both parasitic worms from the class Trematoda, and they are only distantly related to the triclads (18) within the platyhelminths.

Interestingly, searches for Hox/HOM fragments in the planarian genera *Polycelis* and *Phagocata* were in both cases presumed to be exhaustive—i.e., a large number of independent clones were sequenced (100 for *Phagocata* and 94 for *P. nigra*). Despite this fact, the samples of fragments obtained are not perfectly superimposable. Both species show "orphan" genes (Pnox5 and Pnox6 from *Polycelis*; PWoxB, PWoxF, and PWoxG from *Phagocata*). It seems unlikely that these results reflect a true difference between the sets of the Hox/HOM genes of the two organisms because of the close phylogenetic relationship between these two species. It is more likely due to amplification bias (PCR selection, ref. 27). Another point to remember is that the degenerate primers are designed using the sequence data of *Drosophila* and vertebrate homeoboxes. This result means that we do not know whether they are able to amplify the homeoboxes of all the Hox/HOM genes in a very distant organism. The striking differences between genomic and cDNA PCR may be explained both by the relative frequencies of messengers in regenerating planarian total RNA and by the presence of large introns within some of the homeoboxes, rendering them difficult to amplify from genomic DNA.

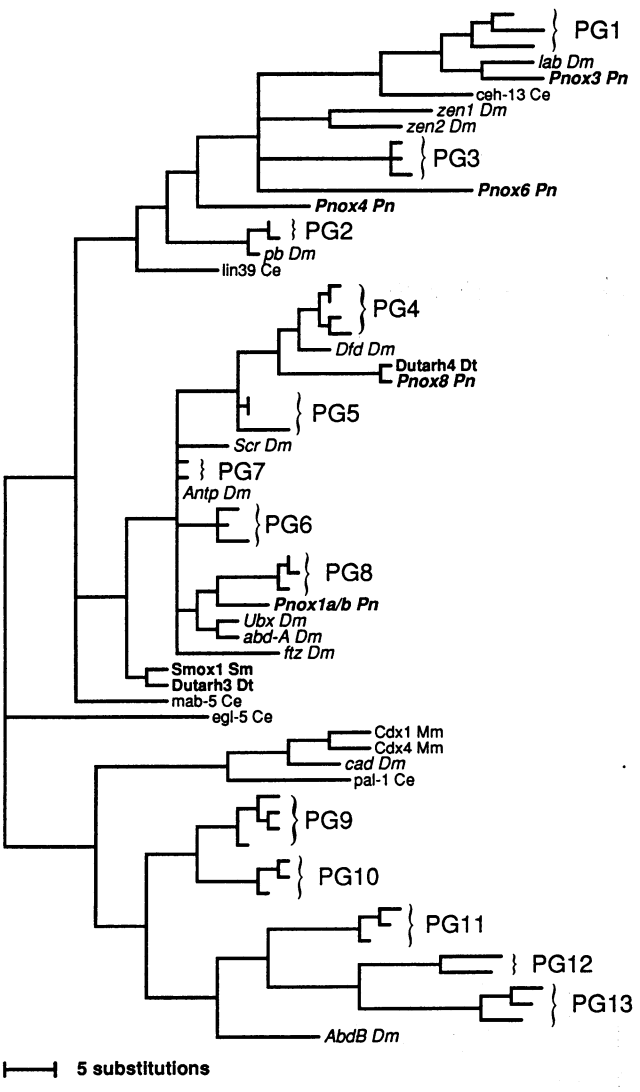

FIG. 3. Majority consensus of the 696 equally most parsimonious trees found by the program PAUP 3.0 with a sample of homeodomain peptide sequences including the seven platyhelminth *Antp*-like complete homeobox sequences known to date (Pn, *P. nigra*; Sm, *Schistosoma mansoni*; Dt, *Dugesia tigrina*), the four Hox/HOM homeodomains of *C. elegans* (Ce), all the known Hox/HOM genes of *Drosophila melanogaster* (Dm), and the 38 human Hox/HOM homeodomains assembled in PG1–13. The samples in the two former animals are very likely to be exhaustive, and hence together they represent a good panel of the Hox/HOM motif variability. In addition, *Drosophila cad* and three *cad* relatives have been added due to their sequence similarity with the Hox/HOM consensus. All groupings shown in this tree are found in >90% of the most parsimonious trees. The *cad* relatives are found clustered with the *Abd-B*-like genes, which indicates a possible origin of this family of nonclustered genes from a duplication within the complex. PG*X*; paralogy group *X*.

Classification of the Planarian Homeodomains. Nearly all homeodomains included in Fig. 3 are complete sequences, allowing us to address the question of gene homology far more convincingly than in previous degenerate PCR work.

The gene Pnox3 is positioned very close to the *Drosophila* gene labial (49/60 amino acid identities, Fig. 4) and even closer than are the three human labial homologues (HOXA1, HOXB1, HOXD1; Fig. 3). Interestingly, another group of PCR fragments related to labial has been amplified from all three species studied in this work (Pnox2, Pfox2, and Dlox2). They are very similar to Pnox3 and may be related to the PCR fragment PWoxA from *Phagocata.*

The Pnox1a and Pnox1b homeodomains are identical in peptide sequences but differ markedly at the nucleotide level (24% difference). In the parsimony tree, Pnox1a/b is found clustered with *Ubx, abd-A*, and the PG8 of vertebrates. This assignation of homology is further supported by some specific similarities in the N-terminal flanking sequence of the homeodomain between Pnox1a/b and *abd-A*, including a conserved splice site position.

Two other full-length homeoboxes have been obtained that are more difficult to classify: Pnox4 is most similar to the genes of the PG3 from vertebrates. It is clustered with the inferred ancestral gene of this group in a distance tree. In a parsimony analysis, this position depends solely on Val-14, which is specific to these vertebrate homeodomains. We cannot, therefore, claim that Pnox4 is the orthologue of the PG3 genes from vertebrates. However, we know from a PCR screen in *Limulus* (28), a chelicerate, that a separate PG3 ancestor gene was already in the cluster of the coelomate ancestor. The position of Pnox6 is even less clear. Its homeodomain shows very little similarity to any other. It is found branched within a large Hox/HOM cluster in a general NJ tree of homeodomains (data not shown). Its assignation to a large group comprising *lab* and PG3 homologs in Fig. 3 is probably a "long branch" artifact, as the similarity of Pnox6 with these genes is very low.

Concerning the short degenerate PCR fragments: Pnox7, one of the short cDNA sequences, is probably orthologous to Dutarh3, a homeobox isolated from *Dugesia tigrina* (26), which branches as an outgroup of a large *Antp/Dfd* class in our parsimony analysis. These two genes may be members of either the *Antp* class or the *Dfd* class, but the homeobox sequence data are not sufficient to discriminate between these two origins. Pnox8, another cDNA clone, is very similar to Dutarh4, which branches as a relative of the *Dfd* class in our tree. Pnox5 is very remote from any other known sequence, so that we cannot claim it represents a Hox/HOM gene as defined earlier.

The last of the cDNA clones recovered, PnMox, is very similar to the *Mox* family genes (29), which are mesoderm-patterning genes not found clustered in homeotic complexes.

Evolutionary Hypotheses. The zootype concept (30) states that a synapomorphy of the Metazoa is the possession of a series of related transcriptional regulators specifying position along the anteroposterior body axis. Data revealing homologous genes in arthropods and chordates have allowed the inference of the presence of these genes with identical functions in the ancestor of the coelomate animals. Slack *et al.* (30) have gone further and have predicted the presence of at least six Hox/HOM genes, as well as others in the ancestor of all metazoa. To test this strong hypothesis, we have looked for evidence of these ancestral Hox/HOM genes in what is generally presumed to be the earliest branch of the triploblastic metazoa. Analysis of the genes found in platyhelminths and of their expression patterns should allow the state of the ancestral triploblastic Hox assemblage to be inferred.

The detailed analyses presented here of the protein sequences of the Antennapedia-like genes found to date in platyhelminths by us and by other groups provide compelling evidence for the existence of homologues of genes of the *lab, Dfd*, and *Antp* classes, and possibly of the *pb/PG3* class, suggesting that these classes originated in an ancestor of the triploblasts. Most probably, Hox genes already had a role in patterning the "head" and "trunk" of the common ancestor of the triploblasts.

The similarity of Pnox3 to its probable homologues in coelomates (*Drosophila* labial and the vertebrate PG1 genes) is very striking (Fig. 4). *Ceh13*, the *C. elegans* labial homologue, has a much more divergent homeodomain. The same is true for the presumed *Antp*-class genes (Pnox1a/b and Smox1 for the platyhelminths, *lin39* and *mab5* for *C. elegans*), although the platyhelminths are traditionally assumed to be more distant from the coelomates than is *C. elegans.* This result could indicate either that nematodes are a deeper branch of the metazoan phylogenetic tree than usually assumed or that the rate of evolution of these genes has been accelerated in nematodes. The apparent absence of *Abd-B*-like genes in the platyhelminths, although not conclusive, may indicate that they

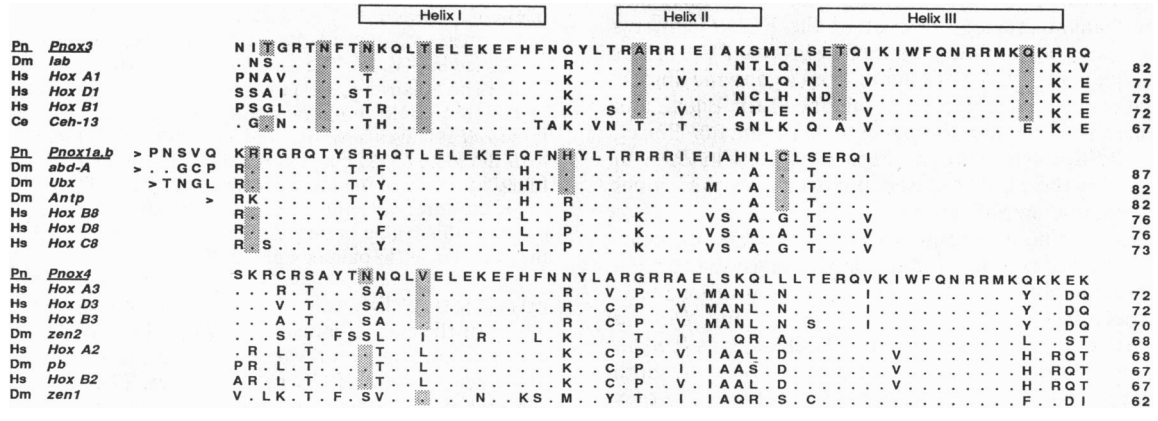

FIG. 4. Alignments of the putative amino acid sequences of the homeodomains of the *P. nigra* genes Pnox1a, Pnox3, and Pnox4 with the most similar homeodomains from various metazoans. The species represented are as follows: Dm, *Drosophila melanogaster*; Ce, *C. elegans*; Hs, *Homo sapiens*; and Pn, *P. nigra.* Global levels of similarity between planarian homeodomains and other sequences are indicated at right (in percentage of identical amino acids). Positions of putative splice acceptor sites upstream to the homeobox are indicated by arrowheads in Pnox1a and related sequences. The shaded residues represent "reliable identities," which allow grouping of the planarian homeodomains with the other sequences in the parsimony analysis. This grouping shows the extensive similarity of Pnox3 with *Drosophila* labial and labial-like genes. Only two residues support the alliance of Pnox4 with the *pb*-like or the *Hox3*-like genes in the PAUP analysis. The global similarity of Pnox4 with these genes is, however, low. The suggested origin of Pnox4 from a *pb/Hox3*-like ancestral gene is therefore speculative. Pnox1a/b are most similar to *Ubx* and *abd-A* in global similarity. Together with these *Drosophila* homeodomains, it branches with PG8 genes of vertebrates.

did not exist in the ancestral triploblast, and this is supported to an extent by their apparent absence in cnidarians. Only searches in further phyla can make this hypothesis more credible.

The "orphan" genes identified here, having no clear homologues elsewhere (Pnox5, Pnox6, and, to a lesser extent, Pnox4) might be Hox/HOM genes, strictly homologous to coelomate genes but with so derived sequences as to render identification of homology impossible. In this case, homology would have to be deduced using functional criteria such as expression patterns. More plausibly, the orphan genes might have arisen as duplications of some of the original genes in the cluster and have diverged rapidly in sequence from their parental gene. This process of duplication has certainly given rise to the multitude of *Abd-B*-like genes in the vertebrate clusters. Some of these duplications have also given rise to various Antennapedia-like genes in flies now co-opted from their original role in position determination (*ftz*, *zen1*, and *zen2*).

The present data from the platyhelminths prove the presence of an elaborate family of Hox/HOM genes [if we consider the groups of orthology identified above, add the "orphan" genes identified in *Polycelis, Phagocata*, or *Dugesia*, and consider the pairs of related fragments (Pnox1a and Pnox1b, Pnox2, and Pnox3; PWoxF and PWoxG) as recent duplications, we conclude that the ancestor of the triclad planarians already had a family of *at least* seven Hox/HOM genes]. This confounds the "gradist" prediction that animals emerging earlier in evolution than the coelomates (traditionally called "lower" metazoans) have a simpler homeotic cluster, if any, than the "higher" metazoans. The data available do not allow us to predict the distribution of these genes in one or more possible clusters, if existing, but the likely tetraploid ancestral condition of *P. nigra* and the presence of two pairs of closely related sequences might suggest the presence of two clusters in this species.

Only extensive work on the expression of these genes, as well as a chromosomal mapping of a potential homeotic cluster, will allow confirmation of the designations of homology given to the genes presented here, thus allowing a rigorous test of the zootype concept.

We thank André Adoutte for support throughout this work; Michael Akam and Michalis Averof for welcoming G. Balavoine in Wellcome/Cancer Research Campaign Institute; and Marie-Josèphe Monnot, Hervé Philippe, Vincent Colot, and Cécile Chaudat for their help. This work was supported by the Centre National de la Recherche Scientifique, the Université Paris-Sud, and by a short-term fellowship from the European Molecular Biology Organization. M.J.T. was supported by a "poste rouge" from the Centre National de la Recherche Scientifique.

1. Bürglin, T. R. (1994) in *Guidebook to the Homeobox Genes*, ed. Duboule, D. (Oxford Univ. Press, Oxford, U.K.), pp. 27–71.
2. Krumlauf, R. (1994) *Cell* **78,** 191–201.
3. Averof, M. & Akam, M. (1993) *Curr. Biol.* **3,** 73–78.
4. Shankland, M. (1994) *BioEssays* **16,** 801–808.
5. Pendleton, J. W., Nagai, B. K., Murtha, M. T. & Ruddle, F. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 6300–6304.
6. Bürglin, T. R. & Ruvkun, G. (1993) *Curr. Opin. Gen. Dev.* **3,** 615–620.
7. Webster, P. & Mansour, T. (1992) *Mech. Dev.* **38,** 25–32.
8. Bartels, J. L., Murtha, M. T. & Ruddle, F. (1993) *Mol. Phyl. Evol.* **2,** 143–151.
9. Oliver, G., Vispo, M., Mailhos, A., Martinez, C., Sosa-Pineda, B., Fielitz, W. & Ehrlich, R. (1992) *Gene* **121,** 337–342.
10. Schummer, M., Scheurlen, I., Schaller, C. & Galliot, B. (1992) *EMBO J.* **11,** 1815–1823.
11. Murtha, M., Dick, M., Ruddle, F. H. & Buss, L. (1991) *J. Exp. Zool.* **260,** 413–416.
12. Naito, M., Ishiguro, H., Fujisawa, T. & Kurosawa, Y. (1993) *FEBS Lett.* **333,** 271–274.
13. McGinnis, W. & Krumlauf, R. (1992) *Cell* **68,** 283–302.
14. Kenyon, C. (1994) *Cell* **78,** 175–180.
15. Schubert, F. R., Nieselt-Struwe, K. & Gruss, P. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 143–147.
16. Hyman, L. H. (1951) *The Invertebrates: Platyhelminthes and Rhynchocoela* (McGraw–Hill, New York).
17. Adoutte, A. & Philippe, H. (1993) in *Comparative Molecular Neurobiology*, ed. Pichon, Y. (Birkhäuser, Basel), pp. 1–30.
18. Riutort, M., Field, K. G., Raff, R. A. & Baguña, J. (1993) *Biochem. Syst. Ecol.* **21,** 71–77.
19. Philippe, H., Chenuil, A. & Adoutte, A. (1994) *Development (Cambridge, U.K.)* Suppl., 15–25.
20. Chomczynski, P. & Sacchi, N. (1987) *Anal. Biochem.* **162,** 156–159.
21. Murtha, M., Leckman, J. & Ruddle, F. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 10711–10715.
22. Philippe, H. (1993) *Nucleic Acids Res.* **22,** 5264–5272.
23. Swofford, D. L. (1990) User's Documentation for PAUP, Phylogenetic Analysis Using Parsimony (Ill. Nat. Hist. Surv., Champaign), Version 3.0.
24. Swofford, D. L. & Olsen, G. J. (1990) in *Molecular Systematics*, eds. Hillis, D. M. & Moritz, C. (Sinauer, Sunderland, MA), pp. 411–501.
25. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4,** 406–425.
26. Tarabykin, V. S., Lukyanov, K. A., Potapov, V. K. & Lukyanov, S. A. (1995) *Gene* **158,** 197–202.
27. Wagner, A., Blackstone, N., Bartels, J., Cartwright, P., Dick, M., Misof, B., Murtha, M., Pendleton, J., Snow, P. & Wagner, G. (1994) *Syst. Biol.* **43,** 250–261.
28. Schierwater, B., Cartwright, P., Dick, M. & Buss, L. W. (1993) *Mol. Phyl. Evol.* **2,** 185–192.
29. Candia, A. F., Hu, J., Crosby, J., Lalley, P., Noden, D., Nadeau, J. H. & Wright, C. V. E. (1992) *Development (Cambridge, U.K.)* **116,** 1123–1136.
30. Slack, J. M., Holland, P. W. H. & Graham, C. F. (1993) *Nature (London)* **361,** 490–492.