Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 7067–7071, September 1989
Evolution

Duplication of large genomic regions during the evolution of vertebrate homeobox genes

(cladistic analysis/gene clusters/gene families)

KLAUS SCHUGHART*, CLAUDIA KAPPEN*, AND FRANK H. RUDDLE*†

Departments of *Biology and †Human Genetics, Yale University, New Haven, CT 06511

Contributed by Frank H. Ruddle, June 19, 1989

ABSTRACT **The phylogenetic relationships of 21 murine *Antp*-class (*Drosophila* mutation *Antennapedia*-type class) homeobox genes have been analyzed, and several groups of related genes have been identified. The murine *Antp*-class homeobox genes are localized within four gene clusters. The similar structural organization of the four gene clusters strongly suggests that genes within a group of related *Antp*-class homeobox genes are derived from duplications of large genomic regions. After the duplication, the gross structures of the homeobox gene clusters have been maintained over a long period of evolutionary time, indicating that the specific organization of genes within a cluster may be of functional importance.**

The analysis of gene families has contributed to our understanding of the molecular evolution of genes and genome structures. Members of the homeobox gene family are characterized by the presence of a conserved 183-base-pair (bp) nucleotide sequence, the so called homeobox (1, 2). The homeobox was originally described in several *Drosophila* genes involved in pattern formation control during early embryonic development (reviewed in refs. 3 and 4). Homeobox genes have been found in many species (1, 5). At least 26 homeobox genes have been identified in the mouse genome. Most of the murine homeobox genes are organized in four gene clusters (*Hox-1*, *-2*, *-3*, and *-4/-5*; see Fig. 4) on chromosomes 6, 11, 15, and 2, respectively (see refs. 6–9). The 21 murine homeobox genes localized within the four gene clusters share nucleotide sequence similarities of 54–92% to each other and are referred to as the *Antp*-class (*Drosophila* mutation *Antennapedia*-type class) homeobox genes. The high sequence similarity of the homeobox as well as the conservation of the structure and organization of genes within the clusters make the *Antp*-class of homeobox genes an excellent system to study phylogenetic relationships of genes within a gene family.

Similarities between homeobox sequences have been used by several authors to define groups of related murine homeobox genes (6–8, 10–14). In these analyses pairwise comparisons of the amino acid sequences of a limited number of genes were performed. To study the phylogenetic relationships of murine *Antp*-class homeobox genes in greater detail, we performed cladistic analysis of the homeobox sequences and sequence comparisons of regions outside the homeobox from all known murine *Antp*-class homeobox genes. Our results strongly indicate that in the vertebrate lineage, the four homeobox gene clusters have arisen through the duplication of an ancestral gene cluster.

MATERIALS AND METHODS

Nucleotide and amino acid sequences were taken from the literature, except for the *Hox-1.2* cDNA sequence (J. Garbern, personal communication). Trees relating homeobox sequences were constructed by using the PAUP program version 2.4.0 (15). "Non-informative" characters were omitted. Characters are considered as informative only if at least two characters each occur in more than one taxon (15). The SWAP=GLOBAL and MULPARS options in PAUP were used to find the most parsimonious trees. Distance matrices from pairwise comparisons were made by a program kindly provided by C. Stephens (16). No gaps were introduced in the homeobox sequences when analyzed by PAUP or in distance matrices. Dot matrix comparisons were performed with the DNA INSPECTOR IIe program (Textco, W. Lebanon, NH). A list of sequences and references as well as distance matrices and dot-plot comparisons are available on request.

RESULTS AND DISCUSSION

Murine *Antp*-Class Homeobox Genes Belong to Separate Cognate Groups. We have constructed trees relating the nucleotide sequences of the highly conserved homeobox region from 21 murine *Antp*-class homeobox genes using the PAUP computer program (Fig. 1). PAUP is designed for inferring phylogenies under the principle of maximum parsimony (15). Briefly, the program works by adding taxa to a branching cladogram at positions and with branch lengths representing relatedness of sequences. Branch lengths are calculated as the minimal number of steps (transformations from one character state to another) required to explain the observed differences in character states between taxa. Different search options like local and global branch swapping are then used to increase the likelihood of finding the shortest tree. The tree length is calculated as the total number of steps required to explain the occurrence of character states of all taxa in the tree. Fig. 1 represents four equally parsimonious trees created by PAUP that relate the nucleotide sequences of homeobox regions from the *Antp*-class homeobox genes. Several groups of related homeobox sequences (cognate groups) are apparent: the *Hox-2.1* group, the *Hox-2.4* group, the *Hox-2.5* group, the *Hox-1.4* group, and the *Hox-2.7* group. The same sequence relationships, except for the branching order of the *Hox-2.5/-5.3* clade, have been established in trees that were based on the analysis of distance matrices (17).

Also, regions outside the homeobox have been compared from those homeobox genes for which cDNA sequences are known. Dot matrix comparisons of nucleotide and amino acid sequences showed that the *Hox-2.4/-3.1* genes, the *Hox-1.3/-2.1* genes, the *Hox-1.1/-2.3* genes, and the *Hox-1.4/-2.6/-5.1* genes represent groups of closely related homeobox genes (refs. 10, 11, and 18 and references therein; K.S., unpublished results). In the PAUP analysis, the *Hox-1.3/-2.1* genes were also recognized as a group of related sequences. In addition, the comparison of cDNA sequences could identify the *Hox-1.1* and *-2.3* genes as a separate cognate group within the *Hox-2.2/-2.3* group. The similarity of the *Hox-2.6* and *-5.1* sequences outside the homeobox

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.

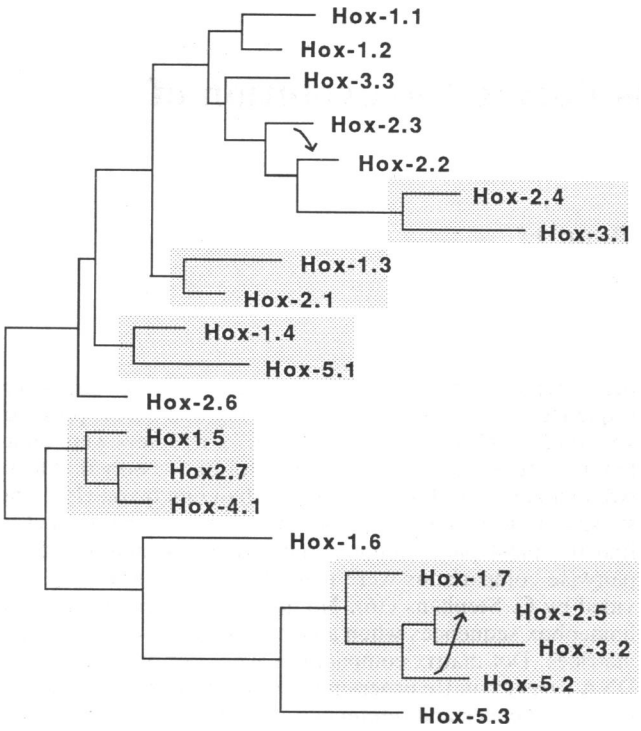

FIG. 1. Tree relating nucleotide sequences of the homeobox regions from the murine *Antp*-class homeobox genes. The tree was created by PAUP. Characters were treated as unordered characters. The tree has been identified by PAUP as one of four equally parsimonious trees. The arrows indicate the branching orders in the other trees. Horizontal branch lengths indicate relative distances of branch points, whereas vertical branch lengths are arbitrary. The tree is 480 steps long and shows a consistency index of 0.431.

region suggests that the *Hox-2.6* gene might represent a member of the *Hox-1.4/-5.1* group.

Partial cDNA sequences have been published for several other homeobox genes allowing us to compare a region between the homeobox and a conserved hexapeptide upstream of the homeobox. Within this region, a splice site has been predicted for all mouse *Antp*-class homeobox genes analyzed. The position of splice sites, the distance between the homeobox, and the hexapeptide are identical for groups of sequences (Fig. 2). Also, the nucleotide and amino acid sequences are highly similar for the same groups. Therefore, these features provide additional criteria to identify related genes within the homeobox gene family. The comparisons of these regions confirmed the results on the relationships of genes in the *Hox-2.1*, *-2.4*, and *-2.5* groups as obtained in the PAUP analysis and defined two additional groups of related homeobox sequences, *Hox-2.2* (*Hox-1.2*, *-2.2*, *-3.3*) and *Hox-2.3* (*Hox-1.1*, *-2.3*) within the closely related genes of the *Hox-2.2/-2.3* group. It is also apparent that the regions between the homeobox and the conserved hexapeptide in the *Hox-2.6* and *-5.1* genes are highly similar. Therefore, we concluded that the *Hox-2.6/-1.4* and *-5.1* genes form a group of related homeobox genes. Changing the branching order of the *Hox-2.6* homeobox sequence in the nucleotide tree in such a way that the *Hox-2.6/-1.4/-5.1* genes are derived from a common ancestor increases the length of the resulting tree by only two steps.

In conclusion, our analyses identified at least seven groups of related homeobox genes in the mouse: *Hox-2.1*, *-2.2*, *-2.3*, *-2.4*, *-2.5*, *-2.6*, and *-2.7*. The *Hox-1.6* and *Hox-5.3* homeoboxes probably represent genes from additional groups from which other members have not yet been described. Fig. 3 represents a phylogenetic tree that not only reflects sequence similarities of the homeobox regions but also structural similarities of regions outside the homeobox. In this tree, the *Hox-2.6* homeobox is connected to the *Hox-1.4/-5.1* clade and the *Hox-2.2/-2.3* group is divided into two separate

```
CGGATGAATTCCTGTGCGG|GTGCCGTGTATGGGAGTCAC  Hox-1.2
 R  M  N  S  C  A |G  A  V  Y  G  S  H
CGCATGAATTCGTGCAACA|GTTCCTCTTTTGGGCCCAGC  Hox-2.2
 R  M  N  S  C  N |S  S  S  F  G  P  S
CGAATGAATTCGCACAGTG|GGGTCGGTTACGGAGCGGAC  Hox-3.3
 R  M  N  S  H  S |G  V  G  Y  G  A  D

                  AGTTCAG|GACCCGAC  Hox-1.1
                   S  S |G  P  D
                  AGCTCAG|GGCCTGAC  Hox-2.3
                   S  S |G  P  D

AAGCTGCACATTAGTCAC|GACAATATAGGTGGCCCAGAA  Hox-1.3
 K  L  H  I  S  H |D  N  I  G  G  P  E
AAGCTTCACATCAGCCAC|GAT...ATGACTGGGCCAGAC  Hox-2.1
 K  L  H  I  S  H |D  .  M  T  G  P  D

AAAGTTCACGTGAGCACG|GTAAACCCCAATTACGCCGGCGGGGAG  Hox-2.6
 K  V  H  V  S  T | V  N  P  N  Y  A  G  G  E
AAGGTGCACGTGAATTCG|GCGAACCCCAACTACACCGGCGGGGAG  Hox-5.1
 K  V  H  V  N  S | A  N  P  N  Y  T  G  G  E

                  CCTCAAG|CAGCCGGA  Hox-2.4
                   P  Q |A  A  G
                  CCCCACG|CTCCTGGG  Hox-3.1
                   P  H |A  P  G

A|ATAACCCGGCTGCCAACTGGCTACATGCTCGCTCC  Hox-1.7
 |N  N  P  A  A  N  W  L  H  A  R  S
A|CCAACCCCTCTGCCAACTGGCTGCACGCCCGCTCT  Hox-2.5
 |T  N  P  S  A  N  W  L  H  A  R  S
```

FIG. 2. Alignment of regions between the conserved hexapeptide and the homeobox from different mouse homeobox genes. A gap has been introduced to align the *Hox-2.1* sequence to the *Hox-1.3* sequence. Splice sites are indicated by arrows. (*Hox-1.2* sequence from J. Garbern, personal communication.)

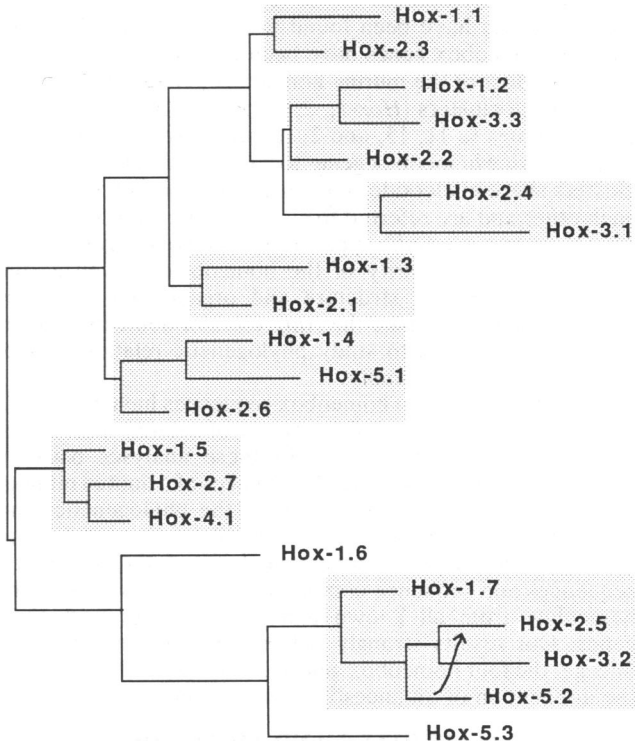

FIG. 3. Phylogenetic tree for nucleotide sequences of the homeobox region from the murine *Antp*-class homeobox genes. The topology of the tree is user-defined (as discussed in the text). Characters were treated as unordered. Horizontal branches correspond to the relative distances of sequences; vertical branch lengths are arbitrary. The arrow indicates the alternative branching of the *Hox-5.2* sequence. The tree is 490 steps long.

cognate groups. The members within each group are predicted to have been derived from a common ancestral gene by gene duplication. As more cDNA sequences become available it should be possible to further test and extend our conclusions about the phylogenetic relationships of murine *Antp*-class homeobox genes, as illustrated in Fig. 3. Our analyses demonstrate that comparisons of regions outside the homeobox need to be included to separate genes with high sequence similarities in the homeobox region into different groups.

When we used only the nucleotide sequences from the homeobox regions in our PAUP analyses, the *Hox-2.2/2.3* groups could not be resolved as two separate groups. This may be due to the extremely high sequence similarities of the homeobox regions from these genes. High similarities of sequences as found in the *Hox-2.2/-2.3* groups can be explained in several ways. For example, genes that have evolved very recently will not have diverged considerably. In such cases, high ratios of transitions versus transversions can generally be expected (19), which we have not observed (data not shown). Therefore, the high sequence similarities are most likely the result of gene conversions or selective constraints for particular sequences.

Murine Homeobox Gene Clusters Have Evolved Through the Duplication of a Large Genomic Region. The analysis of the *Antp*-class homeobox genes has shown that groups of related genes can be identified that are derived from a common ancestor (Fig. 3). When these sequence relationships are projected onto the structrual organization of the genes in the *Hox-1, -2, -3,* and *-4/-5* clusters (Fig. 4), it becomes apparent that the duplication event giving rise to groups of related genes involved the duplication of an entire homeobox gene cluster.

Genes from a particular cognate group can be found at the same position within the *Hox-1* and *Hox-2* clusters (Fig. 4). For example, the *Hox-1.3* homeobox gene is most similar to the *Hox-2.1* homeobox gene, and both genes can be found at the same position within their respective cluster. In addition, the proximal–distal order of *Hox-1* genes (i.e, *Hox-1.2, -1.3,* and *-1.4*) in the *Hox-1* cluster is identical to the order of related genes in the *Hox-2* cluster (*Hox-2.2, -2.1,* and *-2.6*). The same is true for the rest of the gene loci in the *Hox-1* and *Hox-2* cluster. Furthermore, the spacing of cognate genes in different clusters is very similar (Fig. 4), and all genes analyzed so far are transcribed in the same 5′-to-3′ direction. Therefore, genes in the *Hox-1* cluster can be aligned to their cognate genes from the *Hox-2* cluster (Fig. 4), suggesting that both clusters originated from the duplication of a common ancestral gene cluster.

Although the structural analysis of the *Hox-3* cluster has not yet been completed, the genes identified so far can also be aligned to genes from the *Hox-1* and *Hox-2* cluster (Fig. 4), indicating that the *Hox-3* gene cluster might also be derived from a duplication of a precursor cluster. In the *Hox-4/-5* gene cluster four homeobox gene loci have been described to date (refs. 6, 9, and 22). The *Hox-4.1, -5.1* and *-5.2* genes belong to described cognate groups (see above) and are organized in a way similar to genes in other clusters (Fig. 4). These findings suggest that the *Hox-4/-5* cluster also originated from the duplication of an ancestral gene cluster. Consequently, gene loci representing genes related to the *Hox-2.4* to *-2.1* genes (see Fig. 4) would be expected to be present. However, such genes have not been detected by Southern blot hybridizations in the mouse genome (6), whereas at least one more gene (*Hox-5.4*) related to the genes in the *Hox-2.4* group has been described in human (23).

Comparison of homeobox sequences within cognate groups reveal similar degrees of sequence divergence irrespective of their position within the cluster. For example, the *Hox-1.7* and *-2.5* homeoboxes at the 5′ extreme differ in 34 positions, the *Hox-2.2* and *-3.3* homeoboxes in the middle differ in 33 positions, and the *Hox-1.4* and *-2.6* homeobox sequences in the 3′ region of the cluster differ in 32 positions. The same is true for other cognate groups except for genes in the *Hox-2.7* group (see below). These observations show that homeobox sequences from one cluster have diverged from corresponding genes in another cluster to similar extents. The parallel evolution of cognate genes along the entire homeobox gene clusters further supports the idea that the four *Antp*-class homeobox gene clusters derived from the duplication of entire gene clusters.

In conclusion, the comparison of sequence relationships of homeobox genes with the structural organization of genes within clusters strongly suggests that mouse homeobox gene clusters *Hox-1, -2, -3,* and *-4/-5* have evolved in at least two steps. In the first step, an ancestral gene cluster expanded by gene duplications of individual homeobox genes; in the second step, the ancestral gene cluster duplicated several times, and thereby four similarly structured gene clusters were created. The expansion of genes within an ancestral

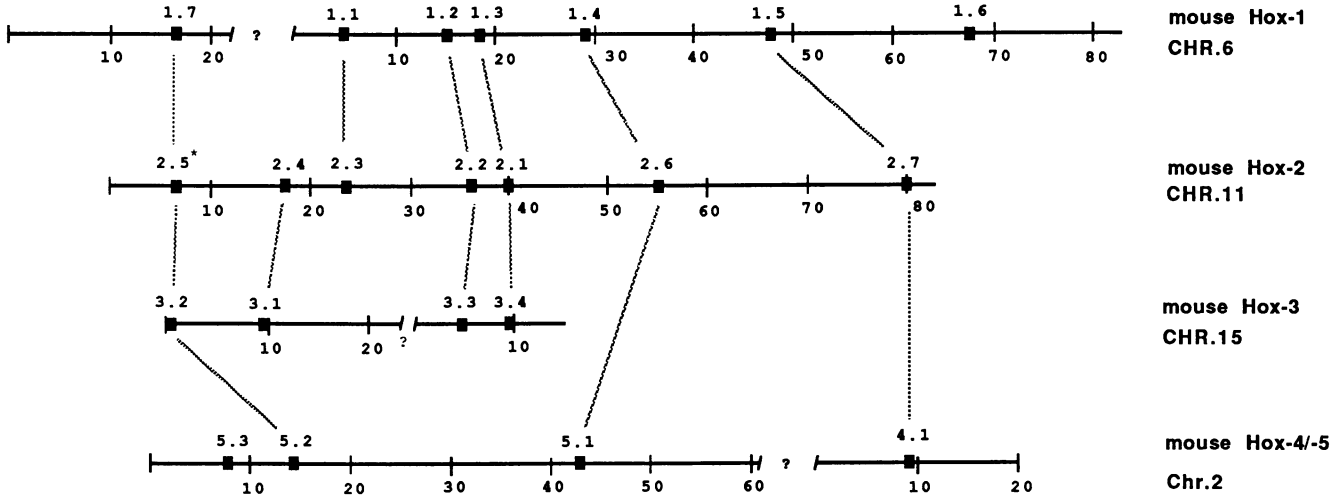

FIG. 4. Structural organization of the murine *Hox-2* gene cluster and presumptive organization of the *Hox-1*, *Hox-3*, and *Hox-4/-5* gene clusters. Vertical bars outline groups of closely related homeobox genes from different gene clusters. Question marks indicate that the precise location and orientation of some homeobox genes within the respective clusters have not yet been determined. The sequence of the murine *Hox-3.4* homeobox (formerly *Hox-6.2*; ref. 20) has not been published. The human *Hox-3.4* homologue (cp11, refs. 17 and 21), however, is a member of the *Hox-2.1* group. The cluster on chromosome 2 has been provisionally named *Hox-4/-5*. For references, see the text.

gene cluster is discussed in a separate paper by Kappen *et al.* (17). The same conclusions can be drawn from the analysis of human *Antp*-class homeobox genes (refs. 17 and 21 and references therein), indicating that the duplication event occurred at least 60–100 million years ago. To our knowledge, the homeobox gene system represents, so far, the best example where duplications of entire gene clusters and subsequent conservation of the organization of genes over such a long period of evolutionary time can be observed.

After the duplication of the gene clusters, strong selection for the conservation of both the sequence and the structural organization of genes within clusters must have existed. This is indicated by the fact that 87% of the nucleotide changes between genes within cognate groups represent silent changes (17) and that the relative order and the spatial arrangement of cognate genes within respective clusters have been conserved (Fig. 4). It should be noted that, to date, no pseudogenes and no large chromosomal rearrangements within clusters have been described. In addition, several authors have noticed a strict correlation between the linear order of genes on the chromosome and anterior expression boundaries of homeobox genes along the anterior–posterior axis of the developing mouse embryo (6, 8, 11, 24–27). These observations together with the conservation of the spatial arrangement of these genes are consistent with the idea that the structural organization of genes within the four clusters is of functional importance.

The spatial arrangement of genes in one cluster is very similar to the spatial arrangement of the corresponding cognate genes in other clusters, but the actual distances are slightly different. For example, the *Hox-1.3*, *-1.4*, and *-1.5* and the *Hox-1.2*, *-2.6*, and *-2.7* genes show a very similar spatial organization in the *Hox-1* and *Hox-2* clusters, respectively. However, distances between genes in the *Hox-2* cluster are somewhat larger (Fig. 4). This might indicate that although protein sequences, direction of transcription, and the relative order of genes have been strongly conserved, regulatory regions in cognate genes have diverged. As a consequence, one would expect that expression boundaries of cognate genes would be very similar but not identical. Such differences in expression patterns have indeed been observed for some cognate genes (*Hox-2.4* and *Hox-3.1*; refs. 8 and 28).

Although we have not found evidence for gene conversions, such events might well have contributed to the evolution of genes in the homeobox gene family. Gene conversions between cognate genes, for example, would result in a smaller number of differences between members of this group when compared with genes in other groups. This is not the case for most cognate groups, where the number of differences are about the same (on the average, 31.7 ± 3.4 changes at the nucleotide sequence level). In the *Hox-2.7* group, however, the number of differences is considerably lower (17.3 positions on the average). This observation could be explained by gene conversions. However, it will be necessary to obtain more sequence data to detect possible gene conversion events. In such cases, it should be possible to find a clustered distribution of changes shared by some members in the *Hox-2.7* group (29).

Duplication of Homeobox Gene Clusters Involved Large Chromosomal Regions. Several groups of genes in the human genome have been described that indicate a paralogous relatedness of chromosomes that carry homeobox gene loci (reviewed in refs. 30 and 31). The same is true for mouse chromosomes carrying homeobox genes (reviewed in ref. 32). These observations strongly suggest that the duplication of homeobox gene clusters was accompanied by the duplication of a large chromosomal region or perhaps the entire chromosome. Duplications of individual chromosomes might be deleterious, and whole genome duplications seem to be more likely (33). Our findings would be consistent with the hypothesis formulated by Ohno (33) that during vertebrate evolution, duplications of the entire genome occurred. We can assume that all four gene clusters might have been present when fish and amphibians arose about 350 million years ago because in *Xenopus* and *Zebrafish*, cognates of homeobox loci from three mouse gene clusters have been described, indicating the presence of at least three gene clusters in these species as well (34–36). The homeobox gene system would represent the most striking example of a remnant of such a genome duplication event.

The homeobox gene system represents an excellent system to study the evolution of vertebrate genomes. The extremely high conservation of both the nucleotide sequence of the homeobox and the structural organization of genes within clusters provides valuable tools to identify homologous sequences in different species and to analyze phylogenetic relationships of genes within this gene family. It will be challenging to investigate whether the expansion of the ancestral gene cluster is accompanied by the appearance of new structural elements and whether the duplication of homeobox gene clusters in vertebrates might have contributed to the establishment of new body plans.

We thank C. Stephens and R. DeSalle for their help with computer programs. We also thank R. DeSalle, M. Goodman, S. Ohno, J. Leckman, and L. Bogarad for their critical discussion of the manuscript. The sequence of the *Hox-1.2* gene has been made available by J. Garbern (National Institutes of Health). The hypotheses outlined in this paper resulted from the work and discussions with many colleagues in the field whose contributions are gratefully acknowledged. This work is supported by National Institutes of Health Grant GM 009966. K.S. and C.K. are supported by the Deutsche Forschungsgemeinschaft (Federal Republic of Germany).

1. McGinnis, W., Garber, R. L., Wirz, J., Kuroiwa, A. & Gehring, W. J. (1984) *Cell* **37**, 403–408.
2. Scott, M. P. & Weiner, A. J. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 4115–4119.
3. Gehring, W. J. (1988) *Science* **236**, 1245–1252.
4. Ingham, P. W. (1988) *Nature (London)* **335**, 25–34.
5. Holland, P. W. & Hogan, B. L. M. (1985) *Nature (London)* **335**, 25–34.
6. Duboule, D. & Dollé, P. (1989) *EMBO J.* **8**, 1497–1505.
7. Schughart, K., Pravtcheva, D., Newman, M. S., Hunihan, L. W., Jiang, Z. & Ruddle, F. H. (1989) *Genomics* **5**, 76–83.
8. Graham, A., Papalopulu, N. & Krumlauf, R. (1989) *Cell* **57**, 367–378.
9. Pravtcheva, D., Newman, M., Hunihan, L., Lonai, P. & Ruddle, F. H. (1989) *Genomics*, in press.
10. Do, M. & Lonai, P. (1988) *Genomics* **3**, 195–200.
11. Graham, A., Papalopulu, N., Lorimer, J., McVey, J. H., Tuddenham, E. G. D. & Krumlauf, R. (1988) *Genes Dev.* **2**, 1424–1438.
12. Hart, C. P., Fainsod, A. & Ruddle, F. H. (1987) *Genomics* **1**, 182–195.
13. Scott, M. P., Tamkun, J. W. & Hartzell, G. W. (1989) *Biochim. Biophys. Acta Rev. Cancer*, in press.
14. Odenwald, W. F., Taylor, C. F., Palmer-Hill, F. J., Friedrich, V., Jr., Tani, M. & Lazzarini, R. A. (1987) *Genes Dev.* **1**, 482–496.
15. Swoffold, D. (1985) *Illinois Natural History Survey* (Champaign, IL).
16. Stephens, C. (1988) HUGHES Human Gene Mapping Library (Yale University, New Haven, CT).
17. Kappen, C., Schughart, K. & Ruddle, F. H. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 5459–5463.
18. Blatt, C., Aberdam, D., Schwartz, R. & Sachs, L. (1988) *EMBO J.* **7**, 4283–4290.
19. Holmquist, R. (1983) *J. Mol. Evol.* **19**, 134–144.
20. Sharpe, P. T., Miller, J. R., Evans, E. P., Burtenshaw, M. D. & Gaunt, S. J. (1988) *Development* **102**, 397–407.
21. Boncinelli, E., Acampora, D., Pannese, M., D'Esposito, M., Somma, R., Gaudino, G., Stornaiuolo, A., Cafiero, M., Faiella, A. & Simeone, A. (1989) *Genome*, in press.

22. Lonai, P., Arman, E., Czosnek, H., Ruddle, F. H. & Blatt, C. (1987) *DNA* **6**, 409–418.
23. Oliver, G., Sidell, N., Fiske, W., Heinzmann, C., Mohandas, T., Sparkes, R. S. & DeRobertis, E. M. (1989) *Genes Dev.* **5**, 641–650.
24. Schughart, K., Utset, M. F., Awgulewitsch, A. & Ruddle, F. H. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5582–5586.
25. Fienberg, A. A., Utset, M. F., Bogarad, L. D., Hart, C. P., Awgulewitsch, A., Ferguson-Smith, A., Fainsod, A., Rabin, M. & Ruddle, F. H. (1987) *Curr. Top. Dev. Biol.* **23**, 233–256.
26. Gaunt, S. J., Sharpe, P. T. & Duboule, D. (1988) *Development*, Suppl. 104, 169–179.
27. Holland, P. W. H. & Hogan, B. L. M. (1988) *Genes Dev.* **2**, 773–782.
28. Utset, M. F., Awgulewitsch, A., Ruddle, F. H. & McGinnis, W. (1987) *Science* **235**, 1379–1382.
29. Eickbush, T. H. & Burke, W. D. (1986) *J. Mol. Biol.* **190**, 357–366.
30. Rabin, M., Ferguson-Smith, A., Hart, C. P. & Ruddle, F. H. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9104–9108.
31. Ruddle, F. H. (1989) in *The Physiology of Growth*, eds. Tanner, J. M. & Priest, M. A. (Cambridge Univ., Cambridge, U.K.), in press.
32. Schughart, K., Kappen, C. & Ruddle, F. H. (1988) *Br. J. Cancer*, Suppl. 9, **58**, 9–13.
33. Ohno, S. (1970) *Evolution by Gene Duplication* (Springer, Heidelberg).
34. Fritz, A. & DeRobertis, E. M. (1988) *Nucleic Acids Res.* **16**, 1453–1469.
35. Fritz, A. F., Cho, K. W. Y., Wright, C. V. E., Jegalian, B. G. & DeRobertis, E. M. (1989) *Dev. Biol.* **131**, 584–588.
36. Njølstad, P. R., Molven, A., Hordvik, I., Apold, J. & Fjose, A. (1988) *Nucleic Acids Res.* **16**, 9097–9111.