刘沛雨 2100012289 信息科学技术学院

Origin and Evolution of Hox Genes

Abstract

Hox genes play a crucial role in the axis patterning of nearly all metazoans. This article explores the origin and evolution of Hox genes, with a focus on the ancestral gene cluster, ProtoHox, and its subsequent divergence into two clusters, Hox and ParaHox. The evolutionary changes in Hox and ParaHox genes, including gene duplications, deletions, and other genomic events, are discussed in relation to the complexity of animal development.

Hox genes consist of the Hox and ParaHox cluster

Hox genes, characterized by the presence of a homeobox, are a large family of homeotic genes that encode proteins with a conserved homeodomain motif. Traditionally, Hox genes were classified based on their DNA sequence, but recent studies have revealed that sequence similarity alone is not sufficient to define Hox genes.

Brooke *et al.* [1] found that a few homeobox genes (such as *Gsx, Xlox,* and *Cdx*) are just as closely related to Hox genes as some of the latter are to themselves. DNA-seq data suggested that the three homeotic genes belong to Hox genes (they are also called Hox-like genes). But actually, they are not in any Hox gene cluster but form another cluster by themselves, which was named ParaHox[1]. Other findings also indicated that in some species of Chordata and Insect, Hox genes do not form an exclusive set (*i.e.,* a cluster independent of other homeotic genes)[2,3]. Molecular phylogenetic analyses manifest that the Hox and ParaHox diverged from a presumptive ancestral gene cluster, ProtoHox. Brooke *et al.* [1] made this inference with data from amphioxus. Later studies confirmed the ProtoHox duplication before the divergence of Cnidaria and Bilateria, which directly caused the emergence of Hox and ParaHox.

Evolution of Hox genes

Clustering of Hox genes is ancient. Coral *Acropora muricata* has a physical linkage of one highly Hox-like gene (*antpC*) to an *Evx* class homeobox gene[4]. Later, D. Chourrout *et al.*[5] reported that two Hox genes, *HoxDb* and *HoxA*, link to an *Evx* class homeobox gene in *Nematostella*, which is consistent with the situation in some mammals. Evidence showed that these Hox-like genes (especially *antpC*) are true Hox genes. So, these physical linkages in Cnidaria suggest that the clustering of Hox genes is ancient, at least for the Hox gene cluster. Now the consensus is that ProtoHox also formed a cluster containing four Hox-like genes in the last common ancestor of the extant metazoans[1].

Composition of primitive Hox and ParaHox gene clusters

The presumptive ProtoHox is the evolutionary ancestor of Hox and ParaHox. The composition of ProtoHox is a question. In their hypothesis, Brooke *et al.* assumed that ProtoHox contains four Hox-like genes (*post*, posterior; *mid*, middle; *X3*, *Xlox-Hox3*; *ant*, anterior), plus *Evx* (**Fig. 1**). Most genes in the primitive Hox cluster have their counterparts in ParaHox, as they are derived from the same ProtoHox gene[1] (*e.g.*, *Gsx* and *Hox1-2* derived from *ant*).

But there are remaining problems in Cnidaria. First, *Xlox* and *Hox3* are lost in almost all extant cnidarians (**Fig. 1**), which points to two possible explanations. *Xlox* and *Hox3* may be lost during the evolution of Cnidaria, or the *X3* is never a part of ProtoHox (*i.e., Xlox* and *Hox3* may originate independently in Bilateria). Molecular phylogenetics indicates that *Hox3* may have been present in the ProtoHox gene cluster, predating the origin of the Hox gene cluster[6]. But there is little convincing evidence. Another question is the origin of the middle Hox (*Hox4-8*). In cnidarians, no exact homologous genes of the middle Hox genes have been identified yet (**Fig. 1**). These questions make the composition of ProtoHox a mystery.

By sequencing all Hox and ParaHox genes in *Nematostella vectensis* and *Hydra magnipapillata*, Chourrout *et al.*[5] revealed their organization in the genome, which strongly supported that the ProtoHox cluster has only 2 anterior genes (*X3* and *ant*, **Fig. 2**). Under this circumstance, the primitive Hox and ParaHox clusters both only include two genes. After the divergence between Cnidaria and Bilateria, the clusters independently expanded in both taxa. The specific time of the expansion events has not been clarified yet. A possible model is that Cnidaria has never acquired middle and posterior Hox genes. These genes may have evolved in Bilateria independently.
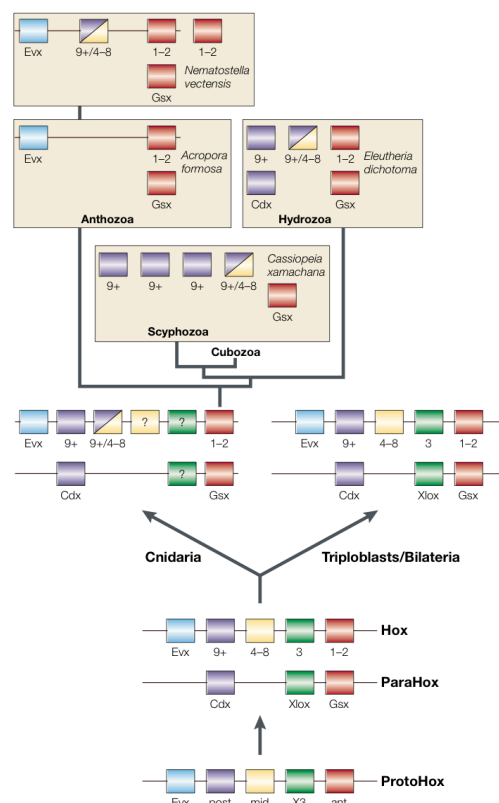

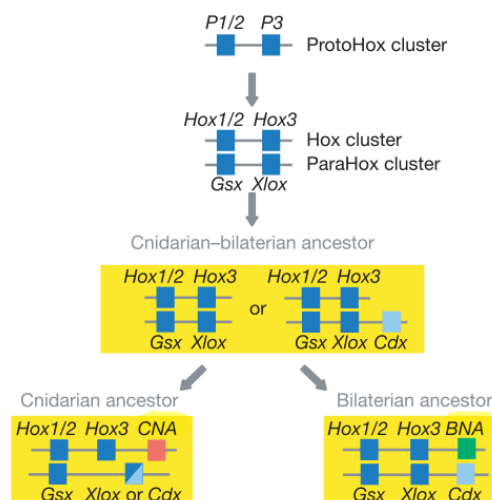
Fig .1 Previous assumption of Hox evolution.[6]

Fig. 2 Updated assumption of Hox evolution.[5]

Evolution of Hox genes

Since the Hox genes were found by Lewis[7] in *Drosophila*, scientists have found similar gene sequences within nearly all species of metazoans. Although Hox genes exhibit extreme conservation in sequence, genomic arrangement, protein structure, and function, they truly have experienced dramatic changes across different animal phyla, especially the Hox gene cluster.
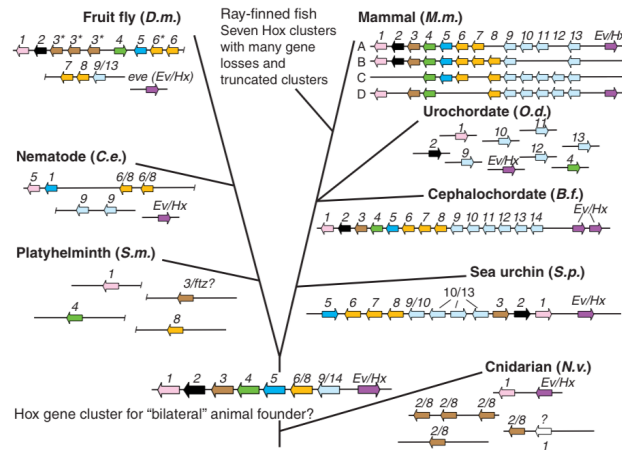

Fig. 3 Brief process and some examples of Hox changes.[13]

In Porifera, the basal group within Metazoan, the presence of true Hox or ParaHox genes is currently uncertain. Even Fragments of Hox-like genes retained from the ProtoHox have not been found yet. However, the fragments were found in another primitive metazoan taxon, Ctenophora[8]. Interestingly, the fragments are highly similar to the middle Hox genes, which are considered absent in Cnidaria. This provides compelling evidence for the separation of Ctenophora and Cnidaria into two distinct phyla.

True Hox and ParaHox genes appeared in Cnidaria and Bilateria. Although the clusters have been widely conserved, their precise gene composition is variable[6]. Besides the absence of middle Hox genes, *Hox3* and *Xlox*, gene duplications and absence are observed in some extant cnidarians. In superior Bilateria, more complex Hox gene changes occurred. *Oikopleura dioica*, a species of Urochordate, has totally fragmented Hox genes (**Fig. 3**). Its genome encodes anterior and posterior Hox genes, none of which are physically linked[9]. It suggested that the collinearity of the Hox genes' physical organization within the genome and their expression on the axes of the embryo is not necessary for Hox genes. Most mammalian genomes encode four Hox gene clusters, A, B, C, and D, each of which shows some gene absence from two or more of the homology groups[10] (**Fig. 3**). Some species of fish also experienced cluster duplications (up to 7 Hox clusters) [11] (**Fig. 3**). These duplication events may correlate with whole genome duplication events (WGD) during animal evolution.

Correlations between the Hox evolution and developmental complication

WGD events provide raw materials for evolution, and more Hox clusters lead to the complication of the Hox gene organization. It may be an evolutionary motivation for the complication of animal morphologies. Hox genes play an important role in embryo axis patterning and development. The number of Hox gene clusters generally reflects the complexity of the body architecture. Take Cnidaria and Bilateria as an example. Molecular morphological evidence supports Cnidaria as the sister group to bilateral animals. Cnidaria has no mesodermal structures, with fewer Hox and ParaHox genes than Bilateria. But this correlation isn't always true. Sea urchin *Strongylocentrotus purpuratus* has a single copy of almost the entire complement of

Hox homologs, but only in one cluster[12]. Its complicated structure arises from a small quantity of Hox genes. Perhaps another mechanism is introduced to regulate the development process.

Conclusions

The clustering of Hox genes dates back to the ancient ProtoHox cluster and has undergone significant evolutionary changes. The divergence of Hox and ParaHox from ProtoHox, along with subsequent gene duplications and deletions, has contributed to the complexity of animal morphogenesis. Understanding the evolution of Hox genes provides valuable insights into the field of *evo-devo*.

References

1. Brooke, N., Garcia-Fernàndez, J. & Holland, P. The ParaHox gene cluster is an evolutionary sister of the Hox gene cluster. *Nature* **392**, 920–922 (1998).

2. Gauchat, D. *et al.* Evolution of Antp-class genes and differential expression of Hydra Hox/ParaHox genes in anterior patterning. *Proc. Natl Acad. Sci USA* **97**, 4493–4498 (2000).

3. Finnerty, J. R. Homeoboxes in sea anemones and other nonbilaterian animals: implications for the evolution of the Hox cluster and the Zootype. *Curr. Top. Dev. Biol.* **40**, 211–254 (1998).

4. Miller, D. J. & Miles, A. Homeobox genes and the Zootype. *Nature* **365**, 215–216 (1993).

5. Chourrout, D., Delsuc, F., Chourrout, P. *et al*. Minimal ProtoHox cluster inferred from bilaterian and cnidarian Hox complements. *Nature* **442**, 684–687 (2006).

6. Ferrier, D., Holland, P. Ancient origin of the Hox gene cluster. *Nat Rev Genet* **2**, 33–38 (2001).

7. Lewis, E. A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565–570 (1978).

8. Finnerty, J. R. *et al*. Homeobox genes in the Ctenophora: identification of paired-type and Hox homologues in the atentaculate ctenophore, Beroë ovata. *Mol. Mar. Biol. Biotech.* **5**, 249–258 (1996).

9. Seo, HC., Edvardsen, R., Maeland, A. *et al*. Hox cluster disintegration with persistent anteroposterior order of expression in *Oikopleura dioica*. *Nature* **431**, 67–71 (2004).

10. Maconochie M, Nonchev S, Morrison A, Krumlauf R. Paralogous Hox genes: function and regulation. *Annu. Rev. Genet.* **30**, 529 (1996).

11. Hurley I, Hale ME, Prince VE., Duplication events and the evolution of segmental identity. *Evol. Dev.* **7**, 556 (2005).

12. R. A. Cameron *et al*., J. Exp. Zoolog. B. Unusual gene order and organization of the

sea urchin hox cluster. *Mol. Dev. Evol.* **306B**, 45 (2006).

13. Derek Lemons, William McGinnis, Genomic Evolution of Hox Gene Clusters. *Science* **313**,1918 (2006).