

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/6785337>

# Genomic Evolution of Hox Gene Clusters

Article in *Science* · October 2006

DOI: 10.1126/science.1132040 · Source: PubMed

---

CITATIONS

302

---

READS

766

2 authors, including:



[William McGinnis](#)

University of California, San Diego

129 PUBLICATIONS 15,084 CITATIONS

SEE PROFILE



## Genomic Evolution of Hox Gene Clusters

Derek Lemons, *et al.*  
*Science* **313**, 1918 (2006);  
DOI: 10.1126/science.1132040

**The following resources related to this article are available online at [www.sciencemag.org](http://www.sciencemag.org) (this information is current as of February 27, 2007):**

**Updated information and services**, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/313/5795/1918>

This article **cites 49 articles**, 20 of which can be accessed for free:

<http://www.sciencemag.org/cgi/content/full/313/5795/1918#otherarticles>

This article has been **cited by** 1 article(s) on the ISI Web of Science.

This article appears in the following **subject collections**:

Evolution

<http://www.sciencemag.org/cgi/collection/evolution>

Information about obtaining **reprints** of this article or about obtaining **permission to reproduce this article** in whole or in part can be found at:

<http://www.sciencemag.org/help/about/permissions.dtl>

# Building the Body from Genes

circadian rhythms, phototaxis, or other light-dependent behaviors. But, if vertebrates are an example, these two photodetection systems functioned together, rather than remaining separate. Although the remaining five families of opsins have not been fully characterized, it seems probable that they also respond to light, and organisms use the information they provide.

## Genomics and Eye Evolution

For decades, scientists have given considerable attention to the primary imaging system in vertebrates, myopically focused on the function of rod and cone photoreceptors and the visual information they deliver. The discovery that animals have multiple parallel pathways to extract information from light and that these coexist in invertebrates, as well as in the eyes of vertebrates, offers new vistas for discovery in development, function, and evolution of eyes and these other novel systems. Genomics could now be used to identify gene regulatory network kernels, similar to those proposed for body plans, for eyes and their parallel systems. Development in a broader phyletic sample of invertebrate eyes could be instructive in helping identify such developmental networks and also for locating other photosensitive systems. Genetic methods have been used to reveal how photoreceptive ganglion cells interact with conventional photoreceptors functionally in mice, and these techniques could now be extended to identify the functions of the other opsin-based systems. Finally, there are abundant evolutionary questions that might be resolved through genomic approaches. Are the inner retinal neu-

rons actually derived from photosensitive precursors? Are there other convergent optical systems like that of cephalopods and vertebrates with common genetic substrates that could be identified and compared? Is the unusual new opsin identified in the parietal eye (9) widespread and will its novel phototransduction system shed light on evolution? Light has been such an important source of information that evolution has exploited it in many ways that remain to be discovered and understood.

## References and Notes

1. C. Darwin, *The Origin of Species by Means of Natural Selection* (John Murray, London, 1859).
2. M. F. Land, D.-E. Nilsson, *Animal Eyes* (Oxford Animal Biology Series, Oxford Univ. Press, Oxford, 2002), pp. 1–15.
3. R. D. Fernald, *Curr. Opin. Neurobiol.* **10**, 444 (2000).
4. BruinFly, [www.bruinfly.ucla.edu](http://www.bruinfly.ucla.edu).
5. J. Chen *et al.*, *PLoS Biol.* **3**, e59 (2005).
6. R. W. Williams, R. C. Strom, G. Zhou, Z. Yan, *Semin. Cell Dev. Biol.* **9**, 249 (1998).
7. A. Packard, *Biol. Rev.* **47**, 241 (1972).
8. A. Ogura, K. Ikeo, T. Gojobori, *Genome Res.* **14**, 1555 (2004).
9. C.-Y. Su *et al.*, *Science* **311**, 1617 (2006).
10. A. Terakita, *Genome Biol.* **6**, 213 (2005).
11. O. Hisatomi, S. Kayada, Y. Aoki, T. Iwasa, F. Tokunaga, *Vision Res.* **34**, 3097 (1994).
12. K. Nordström, T. A. Larsson, D. Larhammar, *Genomics* **83**, 852 (2004).
13. K. S. Dulai, M. von Dornum, J. D. Mollon, D. M. Hunt, *Genome Res.* **9**, 629 (1999).
14. A. D. Briscoe, *Mol. Biol. Evol.* **18**, 2270 (2001).
15. G. H. Jacobs, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 577 (1996).
16. T. C. Spady *et al.*, *Mol. Biol. Evol.* **23**, 1538 (2006).
17. R. C. Fuller, L. J. Fleishman, M. Leal, J. Travis, E. Loew, *J. Comp. Physiol. A Neuroethol. Sens. Neural. Behav. Physiol.* **189**, 609 (2003).
18. T. H. Goldsmith, R. Wehner, *J. Gen. Physiol.* **70**, 453 (1977).
19. D. Osorio, M. Vorobyev, *Proc. Biol. Sci.* **272**, 1745 (2005).

20. J. H. van Hateren, H. P. Snippe, *J. Comp. Physiol. [A]* **192**, 187 (2006).
21. J. L. Spudich, K. H. Jung, in *Handbook of Photosensory Receptors*, W. R. Briggs, J. L. Spudich, Eds. (Wiley, New York, 2005), pp. 1–21.
22. K. Rohde, N. A. Watson, L. A. Chisholm, *Int. J. Parasitol.* **29**, 511 (1999).
23. S. M. Shimeld *et al.*, *Curr. Biol.* **15**, 1684 (2005).
24. G. J. Wistow, J. Piatigorsky, *Annu. Rev. Biochem.* **57**, 479 (1988).
25. G. J. Wistow, J. Piatigorsky, *Gene* **96**, 263 (1990).
26. S. H. Chiou *et al.*, *Biochem. J.* **273**, 295 (1991).
27. S. I. Tomarev, R. D. Zinovieva, B. Chang, N. L. Hawes, *Biochem. Biophys. Res. Commun.* **248**, 684 (1998).
28. H. Janssens, W. J. Gehring, *Dev. Biol.* **207**, 204 (1999).
29. J. Piatigorsky, *Ann. N.Y. Acad. Sci.* **842**, 7 (1998).
30. J. R. True, S. B. Carroll, *Annu. Rev. Cell Dev. Biol.* **18**, 53 (2002).
31. R. H. H. Kröger, M. C. W. Campbell, R. D. Fernald, H.-J. Wagner, *J. Comp. Physiol. A* **184**, 361 (1999).
32. L. V. Salvini-Plawen, E. Mayr, *Evol. Biol.* **10**, 207 (1977).
33. D. E. Nilsson, *Curr. Biol.* **6**, 39 (1996).
34. D. E. Nilsson, S. Pelger, *Proc. Biol. Sci.* **256**, 53 (1994).
35. W. J. Gehring, K. Ikeo, *Trends Genet.* **15**, 371 (1999).
36. G. A. Wray, *Brain Behav. Evol.* **59**, 327 (2002).
37. J. P. Kumar, *Nat. Rev. Genet.* **2**, 846 (2001).
38. X. Li *et al.*, *Nature* **426**, 247 (2003).
39. A. L. Gorman, J. S. McReynolds, *Science* **165**, 309 (1969).
40. D. Arendt, K. Tessmar-Raible, H. Snyman, A. W. Dorresteyn, J. Wittbrodt, *Science* **306**, 869 (2004).
41. D. M. Berson, F. A. Dunn, M. Takao, *Science* **295**, 1070 (2002).
42. D. M. Dacey *et al.*, *Nature* **433**, 749 (2005).
43. D. Arendt, *Int. J. Dev. Biol.* **47**, 563 (2003).
44. I thank A. Fernald, C. Heller, B. Jasny, G. Laurent, H. McAdams, S. McConnell, B. Ruby, M. Scott, and L. Stryer for valuable discussion and comments on previous versions of this manuscript. R.D.F. is supported by the National Institute of Neurological Disorders and Stroke, NIH, NS 34950, the J. Javits Neuroscience Investigator Award.

10.1126/science.1127889

## REVIEW

# Genomic Evolution of Hox Gene Clusters

Derek Lemons and William McGinnis

The family of Hox genes, which number 4 to 48 per genome depending on the animal, control morphologies on the main body axis of nearly all metazoans. The conventional wisdom is that Hox genes are arranged in chromosomal clusters in colinear order with their expression patterns on the body axis. However, recent evidence has shown that Hox gene clusters are fragmented, reduced, or expanded in many animals—findings that correlate with interesting morphological changes in evolution. Hox gene clusters also contain many noncoding RNAs, such as intergenic regulatory transcripts and evolutionarily conserved microRNAs, some of whose developmental functions have recently been explored.

**H**ox genes encode a large family of closely related transcription factors with similar DNA binding preferences. They have not been found in sponges, protozoa, or plants but are present in multiple copies in cnidarians and all bilaterian animals. As a distinct

branch of the homeobox gene superfamily, Hox genes have been a source of fascination since their discovery because of their powerful functions in diversifying morphology on the head-tail axis of animal embryos. This power is revealed by dramatic duplications of head-tail

axial body structures, called homeotic transformations, that can form when one or more of the Hox genes are activated in inappropriate axial positions in developing animals (*1*). The different HOX transcription factors are expressed in distinct, often overlapping, domains on the head-tail body axis of animal embryos (Fig. 1A), and assign different regional fates to these axial domains. As development proceeds, “head” HOX proteins specify the cell arrangements and structures that result in (for example) chewing organs, “thoracic” HOX proteins specify (for example) locomotory organs, and “abdominal” HOX proteins specify (for example) genital and excretory organs. Not surprisingly, extreme homeotic transformations are lethal at early stages of development. Hox genes are also of great interest because there is abundant correlative evidence that changes in Hox expression patterns and protein functions contributed to

Section of Cell and Developmental Biology, University of California San Diego, La Jolla, CA 92093, USA.

\*To whom correspondence should be addressed. E-mail: [wmcginnis@ucsd.edu](mailto:wmcginnis@ucsd.edu)

a variety of small and large morphological changes during animal evolution (2).

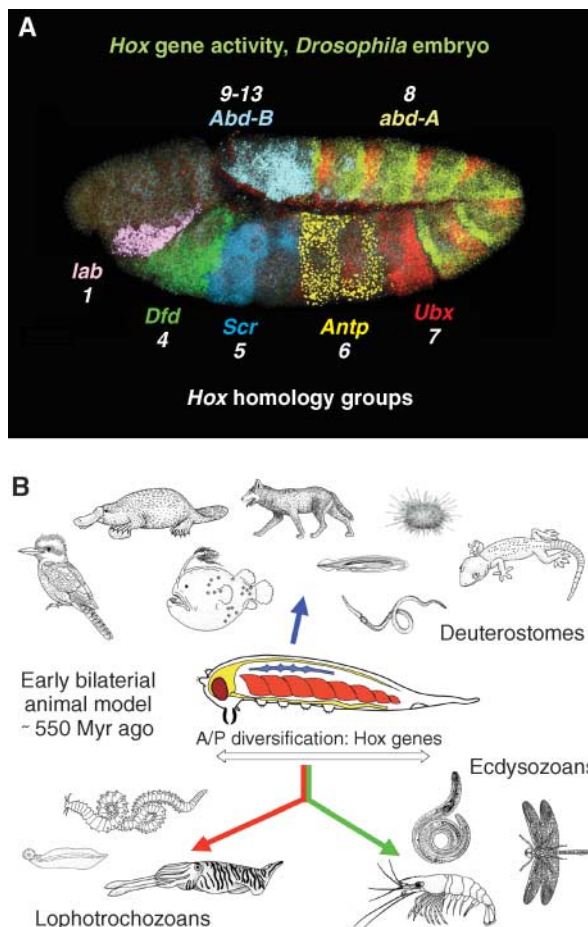
The combination of the aforementioned Hox properties may be the link that relates their functions to the family of MADS box transcription factors that regulate plant developmental patterning. Although structurally unrelated to Hox proteins, the plant MADS box proteins include morphological regulators that have overlapping expression patterns, and similar DNA binding site preferences. Genes encoding plant MADS box proteins can also mutate to provide homeotic transformations of floral structures, and they have been associated with novel expression patterns and functions during plant evolution (3).

### Variations of Colinear Hox Gene Order

Bilateral animals of the deuterostome (e.g., chordates, echinoderms), ecdysozoan (e.g., arthropods, nematodes), and lophotrochozoan groups (e.g., molluscs, annelid worms, and platyhelminth flatworms) are believed to have evolved from a marine, soft-bodied, wormlike ancestor (Fig. 1B). On the basis of the composition of Hox gene clusters in the diverse animals that evolved from this last common bilateral ancestor, this creature possessed a colinear cluster of at least eight Hox-class genes (4, 5), including one of the *Evx* class (Fig. 2). The original *Evx* homolog (*eve*) was identified in *Drosophila* as being required for normal segment number. However, this segmentation function is not conserved in most animals, where *Evx* genes are expressed in extreme posterior regions of developing embryos and serve axial patterning functions similar to those of other Hox genes (6). Hence in Fig. 2, the *Evx* homologs are denoted as *Ev/Hx* to reflect their typical, and presumed ancestral, Hox-like role in tail-region axial patterning.

In some extant animals, the order of Hox genes on the chromosome is roughly colinear with their expression and functional domains on the body axis of embryos (7, 8). The closest to the ideal colinear style of Hox gene arrangement is found on the deuterostome branch in the cephalochordate Amphioxus (*Branchiostoma floridae*) (Fig. 2). The single Amphioxus Hox cluster includes 14 colinear Hox genes closely linked to a pair of *Ev/Hx* genes (9). Both short-range, and very long-range, cis-regulatory elements influence multiple Hox promoters and contribute to maintenance of Hox gene clustering and colinearity

(7, 8). Recent studies in *Drosophila* have provided evidence that supports direct long-range looping contacts between distant regulatory elements, both enhancers and insulators, with specific Hox promoter regions (10, 11).



**Fig. 1.** (A) Confocal image of septuple in situ hybridization exhibiting the spatial expression of Hox gene transcripts in a developing *Drosophila* embryo. Stage 11 germband extended embryo (anterior to the left) is stained for *labial* (*lab*), *Deformed* (*Dfd*), *Sex combs reduced* (*Scr*), *Antennapedia* (*Antp*), *Ultrabithorax* (*Ubx*), *abdominal-A* (*abd-A*), *Abdominal-B* (*Abd-B*). Their orthologous relationships to vertebrate Hox homology groups are indicated below each gene. (B) Illustration with examples of the diversity of body morphologies produced by the expansion of bilateral animals. An artist's conception of the hypothetical last common ancestor of all bilateral animals containing muscle tissue (red), a dispersed "central" nervous system (yellow), blood pumping organ (blue), as well as sensory organs and feeding appendages. This ancestor gave rise to all of the extant animals of the three major bilaterian clades (deuterostomes, ecdysozoans, and lophotrochozoans), which was accompanied by the expansion, diversification, and sometimes simplification, of Hox gene clusters.

However, the idea that most animals preserve colinear clusters of Hox genes, and that this is one of their most important properties, is an oversimplification that is falling under the weight of evidence from increasing numbers of animal genome sequences. An extreme

divergence from colinearity is found in the urochordate *Oikopleura dioica*, whose genome encodes Hox genes of the anterior and posterior groups, none of which are closely linked (Fig. 2) (12). The remaining *Oikopleura* Hox genes are still expressed on the head-tail axis of developing embryos in an order that roughly resembles that of their cephalochordate homologs, although with striking tissue specificities. Another urochordate, *Ciona intestinalis*, shows partial clustering of its Hox gene complement (13) but, like *Oikopleura*, is missing many of the Hox genes from the central homology groups.

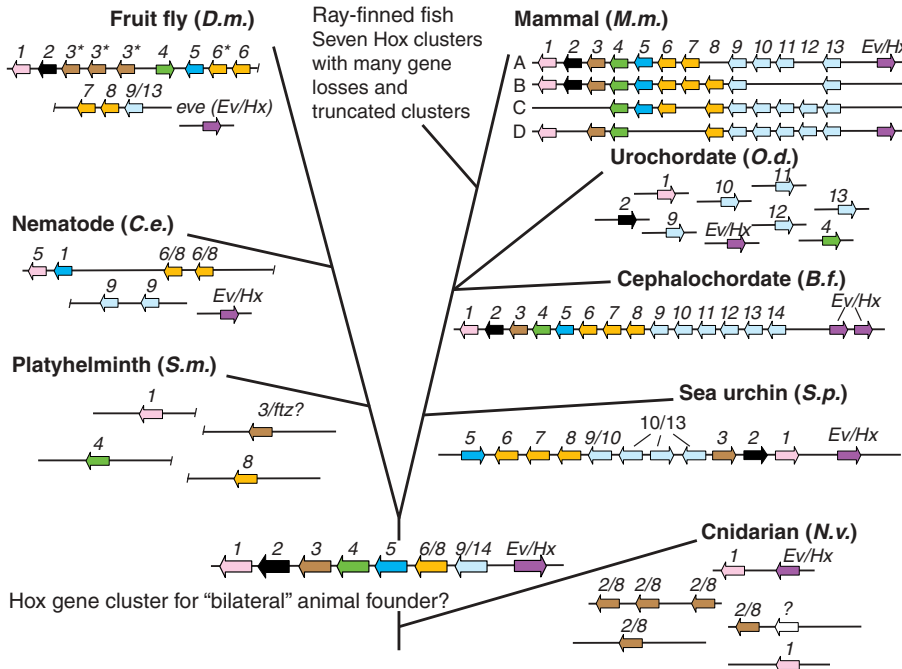
Mammalian genomes encode four Hox gene clusters (14), each of which is missing genes from two or more of the homology groups (Fig. 2), whereas many teleost fishes, having undergone more extensive genomic duplications than their mammalian relatives, possess seven partial Hox clusters (15). Other deuterostomes, like the sea urchin *Strongylocentrotus purpuratus*, have a single copy of almost the entire complement of Hox homologs, but in a scrambled cluster (16).

Among the ecdysozoans, relatively few groups, like nematodes and insects, have had their Hox genomic regions carefully analyzed, although Hox expression patterns are well documented in many arthropods (17). In some insect genomes (e.g., the red flour beetle, *Tribolium castaneum*), a colinear Hox cluster is intact. In other insects, like the fruit fly *Drosophila melanogaster*, the Hox cluster is partially fragmented, and many former Hox genes, now called *zen*, *zen2*, *bcd*, *ftz*, and *eve*, have undergone dramatic changes in protein-coding sequence and expression pattern and have adopted novel developmental patterning functions [e.g., (18); reviewed in (19)]. In another ecdysozoan genome, the nematode *Caenorhabditis elegans*, there has been fairly extensive Hox gene loss and dispersal (20), although almost all the remaining genes still play a role in assigning head-tail axial identities during development.

The lophotrochozoan animals have fascinating morphological differences and diverse Hox genes [e.g., (21, 22)], but the best understood Hox gene arrangement at the genomic level is in the platyhelminth flatworm *Schistosoma mansoni* (23). This parasitic flatworm has only four Hox genes that are dispersed on two different chromosomes and interspersed with many other unrelated genes. The most ancient extant animals with Hox genes are the cnidarians. Whether they



# Building the Body from Genes



**Fig. 2.** Cladogram depicting Hox gene chromosomal organization for representative animals. At the base is shown a cnidarian (*Nematostella vectensis*), which has a dispersed genomic organization of Hox genes and lacks posterior Hox paralogs. The left branch displays fragmented Hox clusters for the lophotrochozoan flatworm *Schistosoma mansoni* and the ecdysozoan fruit fly (*Drosophila melanogaster*) and nematode (*Caenorhabditis elegans*). The right (deuterostome) branch portrays the rearranged but coherent Hox cluster of the sea urchin *Strongylocentrotus purpuratus*, the “prototypical” Hox cluster of *Branchiostoma floridae* (a cephalochordate), the dispersed genomic organization of the Hox genes of a urochordate (*Oikopleura dioica*), and the quadruplicated Hox clusters of a mammal (*Mus musculus*), which remain coherent but have experienced losses of multiple paralogs. Similar to the mammals but not shown diagrammatically, the ray-finned fish have multiple duplicate Hox clusters that are mostly coherent and have experienced gene loss, as exemplified by the zebrafish (*Danio rerio*), pufferfish (*Takifugu rubripes*), and medaka (*Oryzias latipes*). At the base of the cladogram is the likely Hox cluster organization of the last common ancestor of bilaterians (3). Genes are typically assigned to the Hox class if they encode homeodomain sequences that group with the founder HOX protein sequences from *Drosophila* and vertebrate clusters, and then into Hox homology groups arbitrarily designated 1 through 14. Even having a Hox-like homeobox sequence and mapping in a cluster of Hox genes is not an invariably useful standard for Hox axial patterning function in some animals, because one of the *Drosophila* Hox clusters contains the *ftz* (marked 6\* in fly Hox cluster) gene, derived from Hox ancestors, but with novel developmental functions.

have, or their ancestors had, coherent Hox clusters that served oral-aboral (“head-tail”) patterning functions is still hotly debated, although multiple cnidarian species possess homologs of bilaterian Hox genes that are expressed in discrete tissue layers on the oral-aboral axis of developing embryos, and at least two cnidarian species have a *Hox1* homolog closely linked to an *Ev/Hx* homolog in a microcluster (24, 25).

Beginning with Lewis (1), there has been much speculation that changes in Hox gene number may have contributed to morphological evolution. Such ideas are a variation on the old theory that duplicating genes, particularly developmental control genes, may be an initial step in a process that increases regulatory circuit complexity, leading to increased morphological complexity, whereas elimination of control genes may be an initial step leading to less morphological

complexity (26). In one case, gene-replacement experiments showed that two mammalian Hox paralogs have nearly identical functions in a variety of tissues (27), which may be an example of a Hox duplication and functional divergence at an initial stage.

There is currently no rigorous evidence that connects the loss or gain of specific Hox genes or gene complexes with specific morphological changes in different lineages, but there are a number of intriguing correlations. For example, the axial morphology of *Amphioxus* appears relatively simple when compared to the diversity of structures on the head-tail axis of large vertebrates such as fish, reptiles, birds, and mammals (Fig. 1B), which have about four times as many Hox genes as cephalochordates. This apparent increase in Hox regulatory complexity may have contributed to the increased morphological complexity of large vertebrates

(28). In the fish lineage, the teleost (ray-finned) fishes have seven partial Hox clusters, with up to 48 Hox genes in some species, and this is correlated with their great variety in morphology and behavior (15). Conversely, the reduced number of Hox genes in some nematodes is correlated with a relatively simple body architecture compared to many of their ecdysozoan relatives (Fig. 1B). In another ecdysozoan, the crustacean *Sacculina carcini* (a barnacle), a striking reduction in abdominal segments is correlated with the loss of the *abdominal-A* Hox gene (29). A similar argument can be made for the parasitic flatworm, *S. mansoni*, in which reduced Hox regulatory complexity (Fig. 2) may have contributed to its axial architectural simplicity when compared to other lophotrochozoans such as squids (Fig. 1B), which encode at least nine Hox genes in the sepiolid squid *Euprymna scolopes* (30).

On the other hand, adult sea urchins and urochordate tunicates, which exhibit innovative and complex body architectures, have one set of Hox genes, either scrambled or dispersed. Perhaps the novelty of these adult morphologies is dependent on other, equally complex sets of regulatory genes that resemble the Hox genes in their power to diversify morphology but are as yet not well understood.

## Noncoding RNAs of the Hox Gene Clusters

Genetic investigations of the *Drosophila* Bithorax Hox gene cluster (BX-C) have revealed a plethora of important developmental regulatory functions, but genetic and molecular studies identified only three lethal Hox complementation groups corresponding to the HOX protein-coding genes *Hox7* (*Ubx*), *Hox8* (*abd-A*), and *Hox9* (*Abd-B*) in Figs. 2 and 3, (1, 31, 32). Surveys of BX-C transcription have found non-protein-coding transcripts that map to many of the intergenic regions that encode the *bxd*, *iab*, and similar functions (Fig. 3); these regions contain sequences that control levels and patterns of *Ubx*, *abd-A*, and *Abd-B* protein-coding transcripts through a variety of cis and trans mechanisms (33–38).

Chromatin-remodeling proteins play an integral role in epigenetic regulation and maintenance of proper segment-specific transcriptional states of Hox genes (39). Evidence was recently provided for a connection between functions of noncoding Hox cluster transcripts and chromatin-remodeling proteins in Hox regulation (40). ASH1 (a member of the Trithorax group of chromatin-remodeling proteins) was found by in vitro binding assays and chromatin immunoprecipitation (ChIP) analysis to bind transcripts in the *bxd* intergenic region near *Ubx* (40). These intergenic transcripts are normally generated in the same tissues in which *Ubx* is transcribed, and their association with bound ASH1 was correlated with

specific histone methylation patterns characteristic of derepressed chromatin. Depletion of *bxd* region transcripts by small interfering RNA targeting reduced the ChIP association of ASH1 with the locus and was associated with lower levels of *Ubx* transcription. A similar type of regulation may occur in mammalian Hox clusters, because MLL1 (a vertebrate Trithorax group protein) is found to be associated with extensive regions of the human HOXA complex, including much of the intergenic regions (41). These HOXA regions encode many RNAs [as evidenced by EST (expressed sequence tag) clones], some of which, by analogy to *bxd*, might be involved in transcription-dependent chromatin remodeling. However, it remains to be seen if RNA recruitment of ASH1 or other Trithorax group proteins is a curiosity or a more general mode of Hox regulation.

Some of the non-protein-coding transcripts produced in Hox clusters also encode microRNAs (miRNAs). For example, the Hox clusters of many bilaterian animals conserve a sequence for the miRNA miR-10 between their *Hox4* and *Hox5* orthologs (Fig. 3B). In *Drosophila*, miR-10 is predicted to target mRNAs of the neighboring Hox gene *Scr/Hox5* (42), although there are not obvious predicted miR-10 targets in the vertebrate *Hox-5* 3' untranslated region (UTR) sequences.

Arthropod lineages share at least one known miRNA gene, *mir-iab-4*, in the region of the Hox complex containing the abdominal Hox genes (Fig. 3A). Transgene experiments in which a 400-base pair (bp) RNA including the miR-iab-4 hairpin was ectopically expressed in haltere imaginal discs indicate that *Ubx* can be down-regulated by this miRNA in *Drosophila* adult primordia, and patterns of expression of the primary transcript of *iab-4* are complementary to *UBX* protein expression patterns in embryos (43). The relevance of miR-iab-4 to normal developmental patterning awaits the study of mutations that eliminate its function.

At least three of the mammalian Hox clusters encode miRNAs—miR-196a-1, miR-196a-2, and miR-196b—in “abdominal” chromosomal positions similar to those of the *Drosophila iab-4* miRNA, although the miR-196 and miR-iab-4 miRNAs have dissimilar sequences (Fig. 3). A prominent site of expression of miR-196a is in the posterior limb bud, and it has a 1-bp mismatch to a potential target site in

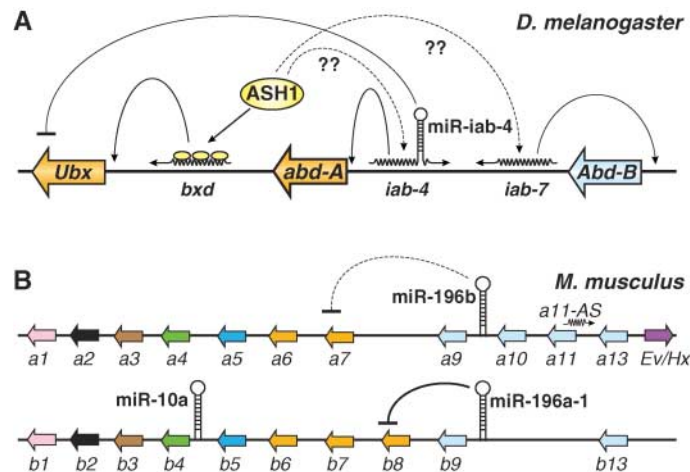
the *Hoxb8* 3' UTR. Analysis of *Hoxb8* transcript cleavage products in the mouse posterior limb bud, and overexpression of a 500-bp RNA containing the miR-196 hairpin in ectopic positions of chick embryos, provide strong evidence that this microRNA cleaves and inactivates *Hoxb8* transcripts (44, 45). All known Hox cluster-encoded miRNAs are expressed in “Hox-like” domains on the head-tail axis of developing embryos (43, 46, 47), so modification of their expression domains, or gain and loss of Hox mRNA target sites, could vary during evolution to influence HOX protein-dependent morphological functions.

initiated from sequences downstream of *Ubx* but extends into the 3' UTR of *Ubx*. Given that the overlapping antisense transcript production precedes *Ubx* protein-coding transcript accumulation, it seems likely that the antisense transcript is involved in inhibition of *Ubx* transcript accumulation (49).

A similar example of apparent antisense-mediated regulation was also found in the mouse *Hoxa* cluster in which transcripts are produced from the opposite strand overlapping the *Hoxa11* gene (50) (Fig. 3). In situ hybridization detecting the antisense transcripts revealed a pattern of expression in the developing limb buds that is complementary to that of *Hoxa11*, suggesting negative regulation by the antisense transcript. It is also possible that the HOXA11 protein represses the antisense transcript, or that a mutually repressive relationship exists.

### Closing Remarks

Genomic analyses have revealed surprising diversity in Hox gene number, organization, and expression patterns in different animals. There are still many animal groups about which little genomic sequence is known, and it remains to be seen how much more variation in Hox gene organization and function will emerge, including the numbers and functions of non-protein-coding RNAs. The property of HOX proteins working as a loosely coordinated system, often with overlapping patterns of expression and function, has apparently fostered their abilities to contribute to morphological change during the evolution of animals. Their colinear arrangement and coordinated regulation in many animals may assist in the maintenance of their overlapping expression patterns. This may have allowed some members of the clusters to subtly and slowly alter their expression patterns and functions to drive groups of cells toward novel structures. But Hox genes still can work as an axial patterning system even when partially dispersed in the genome, and dispersal may foster their rate of functional evolution.



**Fig. 3.** (A) Diagram of the *D. melanogaster* BX-C showing Hox genes (arrows) *Ultrabithorax* (*Ubx*), *abdominal-A* (*abd-A*), and *Abdominal-B* (*Abd-B*), as well as a sampling of noncoding RNAs that derive from the intergenic regions. The intergenic regions *bxd*, *iab-4*, and *iab-7* are transcribed and are known to be involved in proper segment-specific expression of the BX-C Hox genes (interactions with supporting evidence are shown as solid lines). The ability of these regions to affect activation states of the Hox genes is dependent on specific DNA binding sites within these regions (36) and on *ASH1* binding to noncoding transcripts such as RNAs from the *bxd* region, as well as on regulators possibly binding to other BX-C transcripts (dotted lines). (B) Depictions of the mouse *Hoxa* and *Hoxb* clusters. Indicated as hairpins are miRNAs miR-10a, miR-196a-1, and miR-196b along with verified (solid lines) and predicted (dotted lines) interactions with Hox target genes. Also indicated is the position of *Hoxa11* antisense transcripts (*a11-AS*).

Other types of noncoding transcripts, with still undefined genetic functions due to a lack of mutant alleles, have also been discovered in other animal Hox clusters. In a recent survey of Hox gene expression in the centipede *Strigamia maritima*, it was found that transcripts are produced from the 3' region overlapping the centipede *Ubx* gene, antisense to the direction of *Ubx* transcription (48). These transcripts are expressed in the anterior maxillipedal segment, as well as in a subset of neuroectodermal cells in the more posterior limb bearing segments in a pattern complementary to that of *Ubx* transcription. The antisense transcript is

### References and Notes

1. E. B. Lewis, *Nature* **276**, 565 (1978).
2. S. B. Carroll, J. K. Grenier, S. D. Weatherbee, *From DNA to Diversity* (Blackwell Science, London, ed. 2, 2005).
3. V. F. Irish, A. Litt, *Curr. Opin. Genet. Dev.* **15**, 454 (2005).
4. R. de Rosa et al., *Nature* **399**, 772 (1999).
5. J. Garcia-Fernandez, *Nat. Rev. Genet.* **6**, 881 (2005).
6. J. Ahringer, *Genes Dev.* **10**, 1120 (1996).
7. M. Kmita, D. Duboule, *Science* **301**, 331 (2003).

8. J. Deschamps, J. van Nes, *Development* **132**, 2931 (2005).
9. C. Minguiillon *et al.*, *Int. J. Biol. Sci.* **1**, 19 (2005).
10. M. Ronshaugen, M. Levine, *Dev. Cell* **7**, 925 (2004).
11. F. Cleard, Y. Moshkin, F. Karch, R. K. Maeda, *Nat. Genet.* **38**, 931 (2006).
12. H. C. Seo *et al.*, *Nature* **431**, 67 (2004).
13. T. Ikuta, N. Yoshida, N. Satoh, H. Saiga, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 15118 (2004).
14. M. Maconochie, S. Nonchev, A. Morrison, R. Krumlauf, *Annu. Rev. Genet.* **30**, 529 (1996).
15. I. Hurley, M. E. Hale, V. E. Prince, *Evol. Dev.* **7**, 556 (2005).
16. R. A. Cameron *et al.*, *J. Exp. Zool. B Mol. Dev. Evol.* **306B**, 45 (2006).
17. C. L. Hughes, T. C. Kaufman, *Evol. Dev.* **4**, 459 (2002).
18. U. Lohr, L. Pick, *Curr. Biol.* **15**, 643 (2005).
19. C. C. Hsia, W. McGinnis, *Curr. Opin. Genet. Dev.* **13**, 199 (2003).
20. A. Aboobaker, M. Blaxter, *Curr. Opin. Genet. Dev.* **13**, 593 (2003).
21. M. Shankland, E. C. Seaver, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 4434 (2000).
22. V. F. Hinman, E. K. O'Brien, G. S. Richards, B. M. Degnan, *Evol. Dev.* **5**, 508 (2003).
23. R. J. Pierce *et al.*, *Mol. Biol. Evol.* **22**, 2491 (2005).
24. J. R. Finnerty, K. Pang, P. Burton, D. Paulson, M. Q. Martindale, *Science* **304**, 1335 (2004).
25. K. Kamm, B. Schierwater, W. Jakob, S. L. Dellaporta, D. J. Miller, *Curr. Biol.* **16**, 920 (2006).
26. S. Ohno, *Evolution by Gene Duplication* (Springer-Verlag, Berlin, New York, 1970).
27. J. M. Greer, J. Puetz, K. R. Thomas, M. R. Capecchi, *Nature* **403**, 661 (2000).
28. P. W. Holland, J. Garcia-Fernandez, *Dev. Biol.* **173**, 382 (1996).
29. M. Blin, N. Rabet, J. S. Deutsch, E. Mouchel-Vielh, *Dev. Genes Evol.* **213**, 90 (2003).
30. P. Callaerts *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 2088 (2002).
31. E. Sanchez-Herrero, I. Vernos, R. Marco, G. Morata, *Nature* **313**, 108 (1985).
32. M. Regulski *et al.*, *Cell* **43**, 71 (1985).
33. H. D. Lipshitz, D. A. Peattie, D. S. Hogness, *Genes Dev.* **1**, 307 (1987).
34. E. Sanchez-Herrero, M. Akam, *Development* **107**, 321 (1989).
35. E. Bae, V. C. Calhoun, M. Levine, E. B. Lewis, R. A. Drewell, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 16847 (2002).
36. R. A. Drewell, E. Bae, J. Burr, E. B. Lewis, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 16853 (2002).
37. R. K. Maeda, F. Karch, *Development* **133**, 1413 (2006).
38. O. S. Akbari, A. Bousum, E. Bae, R. A. Drewell, *Dev. Biol.* **293**, 294 (2006).
39. L. Ringrose, R. Paro, *Annu. Rev. Genet.* **38**, 413 (2004).
40. T. Sanchez-Elsner, D. Gou, E. Kremmer, F. Sauer, *Science* **311**, 1118 (2006).
41. M. G. Guenther *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 8603 (2005).
42. A. J. Enright *et al.*, *Genome Biol.* **5**, R1 (2003).
43. M. Ronshaugen, F. Biemar, J. Piel, M. Levine, E. C. Lai, *Genes Dev.* **19**, 2947 (2005).
44. E. Hornstein *et al.*, *Nature* **438**, 671 (2005).
45. S. Yekta, I. H. Shih, D. P. Bartel, *Science* **304**, 594 (2004).
46. D. Kosman *et al.*, *Science* **305**, 846 (2004).
47. E. Wienholds *et al.*, *Science* **309**, 310 (2005).
48. C. Brena, A. D. Chipman, A. Minelli, M. Akam, *Evol. Dev.* **8**, 252 (2006).
49. S. I. Grewal, J. C. Rice, *Curr. Opin. Cell Biol.* **16**, 230 (2004).
50. H. M. Hsieh-Li *et al.*, *Development* **121**, 1373 (1995).
51. We thank D. Kosman for the multiplex in situ image in Fig. 1A, A. Veraksa for help with Fig. 1B, and E. Tour for comments on the manuscript. We apologize to the many authors whose papers could not be cited because of space limitations.

10.1126/science.1132040

## REVIEW

# Gene Regulatory Networks in the Evolution and Development of the Heart

Eric N. Olson

The heart, an ancient organ and the first to form and function during embryogenesis, evolved by the addition of new structures and functions to a primitive pump. Heart development is controlled by an evolutionarily conserved network of transcription factors that connect signaling pathways with genes for muscle growth, patterning, and contractility. During evolution, this ancestral gene network was expanded through gene duplication and co-option of additional networks. Mutations in components of the cardiac gene network cause congenital heart disease, the most common human birth defect. The consequences of such mutations reveal the logic of organogenesis and the evolutionary origins of morphological complexity.

The formation of organs and body parts proceeds by sequential gene regulatory steps that dictate cell fates and organize specialized cell types into complex three-dimensional units of structure and function. Studies of heart development and its genetic underpinnings in simple model organisms and in vertebrates have revealed an evolutionarily conserved gene regulatory network consisting of functional interconnections between myogenic transcription factors, their downstream target genes, and upstream signaling pathways that direct cardiac cell fate, myocyte differentiation, and cardiac morphogenesis (1, 2). Comparative genomic analyses of cardiac developmental control genes and their cis-regulatory elements

have also highlighted the conservation of genetic pathways that direct cardiogenesis.

The striking parallels between the transcriptional networks involved in heart development across vast phylogenetic distances support the idea that the evolutionary emergence of hearts with increasing complexity occurred through modification and expansion of an ancestral network of regulatory genes encoding cardiac transcription factors. The expansion of cardiac genetic networks through the duplication of cardiac regulatory genes and the co-option of additional gene networks probably allowed for the addition of new accessory structures, such as chambers, valves, and a conduction system, to a primitive vessel-like heart analogous to that of invertebrates and vertebrate embryos (3). The modular addition of innovations to primitive structures, although speculative, has also been proposed as a mechanism for the genesis of other

vertebrate organs and body structures (4). Insights into the genetic circuits that drive the evolution and development of the heart shed light on general principles of organogenesis and evolutionary origins of morphological complexity, as well as the molecular basis of cardiovascular disease in humans.

## Evolutionary Advancements of the Heart

The most fundamental functional units of all hearts are cardiac muscle cells, which express an array of contractile proteins, such as muscle actin, myosin, troponin, and tropomyosin. The appearance of muscle cells preceded the divergence of Cnidaria (hydra, jellyfish, and corals) and Ctenophora (comb jellies) from the Bilateria, from which mammals descended (~700 million years ago) (Fig. 1) (5). Primordial muscle cells probably resembled the epitheliomuscle cells of Cnidaria and amphioxus, which is thought to be the closest living approximation of the invertebrate ancestor of vertebrates (6–8). These cells probably existed in a primitive gastric pocket where they participated in fluid movement during feeding. Muscle cells in bilaterians are derived from mesoderm, which is believed to have arisen from the gastrodermis of a diploblastic ancestor. The diversification of muscle cells gave rise to skeletal, cardiac, and smooth muscle cells, and further specialization of cardiac muscle cells ultimately yielded atrial and ventricular myocytes, as well as the cells of the mammalian cardiac conduction system.

The first heart-like organ is believed to have appeared over 500 million years ago in an ancestral bilaterian (6, 9–11). It probably resembled the simple tubular vessel-like organs of tunicates and amphioxus, which contain a

Department of Molecular Biology, University of Texas Southwestern Medical Center, 6000 Harry Hines Boulevard, Dallas, TX 75390, USA. E-mail: eric.olson@utsouthwestern.edu