Hox gene evolution in nematodes: novelty conserved
Aziz Aboobaker and Mark Blaxter

The conserved homeobox (Hox) gene cluster is neither conserved nor clustered in the nematode Caenorhabditis elegans. Instead, C. elegans has a reduced and dispersed gene complement that is the result of the loss of Hox genes in stages throughout its evolutionary history. The roles of Hox genes in patterning the nematode body axis are also divergent, although there are tantalising remnants of ancient regulatory systems. Hox patterning also differs greatly between C. elegans and a second ‘model’ nematode, Pristionchus pacificus. The pattern of Hox gene evolution may be indicative of the move to deterministic developmental modes in nematodes.

Addresses
1Institute of cell, Animal and Population biology, Ashworth Laboratories, King’s Buildings, University of Edinburgh, Edinburgh EH9 3JT, UK
2e-mail: mark.blaxter@ed.ac.uk
3e-mail: AzizAboobaker@netscape.net

Introduction
Homeobox (Hox) genes, which are found closely clustered in the genomes of animals, encode homeodomain transcription factors required for patterning their anterior–posterior axis. Comparison of Hox gene complements between phyla suggests that the ancestral protostome had a cluster with at least nine groups of orthologues and that most extant taxa have representatives of all of them [1] (see Figure 1 and [2,3] for recent additions to the Hox ‘zoo’). By contrast, the developmental model organism Caenorhabditis elegans has a reduced and dispersed cluster of Hox genes [4**] that have novel roles and regulation [5**,6]. Although there is no doubt that the Hox gene cluster is a genetic keystone for the bilaterian body plan, for nematodes the breakdown of this cluster may represent a taxon-defining evolutionary change in developmental processes.

Nematodes are renowned for their simple and highly conserved body plan and for having a ‘quasi-invariant’, deterministic, lineage-driven mode of development in comparison to the regulative processes observed in other phyla (for recent descriptions of lineage-mode development in nematodes, see [7–9]). The highly derived Hox gene set of C. elegans was previously interpreted as being representative of ancestral bilaterian features; however, recent evidence suggests that nematodes have evolved from an ancestor in possession of a complete set of Hox genes [4**]. That being the case, we can ask how, why and when did the lineage leading to C. elegans evolve its depauperate and disorganised Hox gene complement?

In this review, we discuss recent data revealing the details of Hox gene function and regulation in C. elegans, comparative analyses of Hox gene functions in other nematodes, and the Hox gene complements of nematodes that are only distantly related to C. elegans.

Evolving the C. elegans Hox cluster: gene loss and fusion
In C. elegans there are only six Hox genes — the anterior gene ceh-13; two distinct central genes, lin-39 and mab-5; and three posterior group genes, egl-5, php-3 and nob-1 — organized in three gene pairs that are spread widely across 5 Mb of chromosome III (Figure 1). In addition, the collinearity between genomic position and regulatory effect has broken down: the anterior gene ceh-13 is positioned downstream of the central gene lin-39. At the sequence level, the C. elegans Hox genes have diverged substantially from their orthologues in other phyla [4**]. Even more revealing is the finding that these six genes represent only four orthologue groups. From this, we can infer that orthologues of Hox2 (also known as Pbh), Hox3, Hox4 (Dfd), Antp, Lox2 (Ubx) and Lox4 (AbdA) have been lost (Figure 2).

Nematoda is a speciose and diverse phylum that has a remarkably conservative body plan incorporating a vermiform body, with only longitudinal muscles; a pseudocoelom, in which the gut and gonads float freely; a ventral nerve cord; and not much that could be called ‘appendages’ (some taxa have well-developed setae and bristles). Thus, the reduced Hox complement in C. elegans (a crown group rhabditid) might be a basic genomic fact in nematodes. In a survey of species across the phylum, however, many ‘extra’ Hox genes have been found,

www.current-opinion.com

Current Opinion in Genetics & Development 2003, 13:593–598

DOI 10.1016/j.gde.2003.10.009

Abbreviations
Hox homeobox gene
MAPK mitogen-activated protein kinase
mes maternal effect sterile
PgC Polycomb group
SAM sterile 5-motif
sop suppressor of pal
TGF-β transforming growth factor-β
tRG tritotax group
VPC vulval precursor cell
wg wingless

Current Opinion in Genetics & Development 2003, 13:593–598
This review comes from a themed issue on Genomes and evolution
Edited by Evan Eichler and Nipam Patel
0959-437X/$ – see front matter © 2003 Elsevier Ltd. All rights reserved.

Current Opinion in Genetics & Development 2003, 13:593–598

www.current-opinion.com

DOI 10.1016/j.gde.2003.10.009
Comparisons of the Hox clusters of flies, worms and humans: the C. elegans `cluster' is depauperate and disorganised. The Hox cluster regions of C. elegans, D. melanogaster, Anopheles gambiae — a mosquito with an unbroken Hox cluster — and Homo sapiens were extracted from Ensembl (http://www.ensembl.org/). Each cluster is mapped at the same scale, with the coloured boxes representing each gene drawn proportionally to the length of the gene on the chromosome. The C. elegans and D. melanogaster clusters are each broken between antennapedia-like genes and more posterior genes, with the two portions separated by ∼4 Mb and ∼10 Mb, respectively; the cluster fragments of both species are still on the same chromosome. The contiguated A. gambiae genome sequence for chromosome 2R does not include AbdB, which is present on an unordered contig of 0.25 Mb known to derive from chromosome 2R. H. sapiens, like other vertebrates, has four clusters, each of which is as compact as the HOXA cluster pictured. Distances are given in megabases.

<table>
<thead>
<tr>
<th>Genome</th>
<th>Chromosome</th>
<th>HOXA</th>
<th>3.9 Mb</th>
<th>9.5 Mb</th>
<th>11.8 Mb</th>
<th>12.4 Mb</th>
<th>12.9 Mb</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. elegans</td>
<td>III</td>
<td>lin-39</td>
<td>ceh-13</td>
<td>mab-5</td>
<td>egl-5</td>
<td>nob-1</td>
<td>php-3</td>
</tr>
<tr>
<td>D. melanogaster</td>
<td>3R</td>
<td>lab</td>
<td>z2</td>
<td>Dfd</td>
<td>ftz</td>
<td>Antp</td>
<td>Ubx</td>
</tr>
<tr>
<td>A. gambiae</td>
<td>2R</td>
<td>lab</td>
<td>pb</td>
<td>zen</td>
<td>Dfd</td>
<td>Scr</td>
<td>ftz</td>
</tr>
<tr>
<td>H. sapiens</td>
<td>7</td>
<td>HOXA</td>
<td>26.6</td>
<td>27.1</td>
<td>62.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comparisons of the Hox clusters of flies, worms and humans: the C. elegans `cluster' is depauperate and disorganised. The Hox cluster regions of C. elegans, D. melanogaster, Anopheles gambiae — a mosquito with an unbroken Hox cluster — and Homo sapiens were extracted from Ensembl (http://www.ensembl.org/). Each cluster is mapped at the same scale, with the coloured boxes representing each gene drawn proportionally to the length of the gene on the chromosome. The C. elegans and D. melanogaster clusters are each broken between antennapedia-like genes and more posterior genes, with the two portions separated by ∼4 Mb and ∼10 Mb, respectively; the cluster fragments of both species are still on the same chromosome. The contiguated A. gambiae genome sequence for chromosome 2R does not include AbdB, which is present on an unordered contig of 0.25 Mb known to derive from chromosome 2R. H. sapiens, like other vertebrates, has four clusters, each of which is as compact as the HOXA cluster pictured. Distances are given in megabases.

Including several that clearly belong to orthologue groups that are absent in C. elegans [4**] (Figure 2).

Although overall Hox gene sequence evolution has been much more rapid in the Nematoda than in other phyla, these new nematode Hox genes also support the attribution of the highly divergent Hox genes of C. elegans into canonical orthologue groups. By mapping the Hox gene complements onto a species phylogeny derived from rRNA, it is also evident that gene loss from the cluster has been piecemeal: in other words, there was not a single, massive loss event that resulted in the C. elegans Hox genotype. Investigation of Hox gene content in a nematode, Paragordius robustus, indicates that this phylum (the closest relative to Nematoda) has a full set of protostome genes (A Aboobaker, M Blaxter, unpublished data). Nematomorphs are all parasitic in arthropods and also have a very simple, vermiciform body plan. Thus, the loss of Hox genes is limited to the nematode lineage and is unlikely to be just the result of the evolution of a simple body plan.

Tantalisingly, this survey also suggests a potential mechanism for Hox gene loss. In the filarial parasitic nematode Brugia malayi, the orthologue of egl-5 is found immediately upstream of an antennapedia-like gene, ant-1 [4**]. The egl-5 and ant-1 genes of B. malayi share a 5' exon encoding the amino-terminal domain of the protein but have very different homeodomain exons. The two mature proteins, classified to different orthologue groups that have existed for perhaps a billion years, derive from a single, alternatively spliced, fused gene. The expansion of the posterior group of genes (to three in C. elegans and two in other nematodes) may be in part compensatory for the loss of central genes.

**Conserved and novel aspects of C. elegans Hox genes**

Have gene loss and cluster breakdown also been accompanied by the loss of functional and regulatory conservation, which is shared by flies and vertebrates, for the remaining C. elegans Hox genes? Hox gene function in C. elegans has some similarity to that in vertebrates and arthropods. Hox genes have roles in specifying cell fates along the body axis, follow spatial collinearity by orthologue group (even if their chromosomal order does not) and can display cross- and auto-regulation. Unlike in other animals, however, Hox gene expression in C. elegans is dependent on lineage rather than position [5*], and only the anterior gene cel-13 and the posterior genes php-3 and nob-1 are required during embryogenesis [10]. Although elimination of both php-3 and nob-1 results in gross posterior defects and posterior-to-anterior cell fate transformations, the AbdB orthologue php-3 seems to have only a minor role in posterior patterning as compared with the role of the more divergent nob-1. The remaining three genes are not essential for embryogenesis, but instead are required for the specification of cell fates and cell migrations in postembryonic development (reviewed in [11]).
As in *Drosophila melanogaster* and other animals, Hox gene functions in *C. elegans* require the nematode homologues of the Hox cofactors *extradenticle* (*ceh-20* and *ceh-40*) and *homothorax* (*unc-62*) [6,12]. Recent studies indicate that additional aspects of function and regulation also may be conserved. In *D. melanogaster*, correct regulation of *labial* (the *ceh-13* orthologue) requires both *wingless* (*wg*), implicating a Wnt pathway, and *decapentaplegic*, implicating a transforming growth factor-β (TGF-β) pathway, as well as an autoregulatory component (recently summarized in [13]).

Streit *et al.* [5^*]* have studied the promoter region of *ceh-13* and have shown, in the nematode, that early embryonic expression may require both Wnt signalling and an autoregulatory loop for maintenance. An element of the *ceh-13* promoter that directs early nematode embryo expression can drive reporter activity in *D. melanogaster* embryos in a *wg*-dependent process. Thus, some regulatory control for *Hox1* orthologues is conserved between deterministic (nematode) and regulative (*fly*) developmental contexts [5^*]*.

Surprisingly, part of the *ceh-13* promoter is also responsible for postembryonic expression in the male tail in a process dependent on the TGF-β pathway, although no patterning defects have been observed in the male tail in *ceh-13* mutants [14]. *ceh-13*, like *mab-5* and *egl-5*, may have been recruited to pattern this sex-specific structure. It is an intriguing possibility that the regulation by a TGF-β
homologue of \textit{labial} in \textit{D. melanogaster} and of \textit{ceh-13} in \textit{C. elegans} may have a common evolutionary origin.

The conserved trithorax group (trG) and Polycomb group (PcG) chromatin-regulating proteins are required for correct global maintenance (trG) and repression (PcG) of Hox gene expression in both vertebrates and flies. Although the \textit{C. elegans} trG protein LIN-59 has been shown to regulate \textit{C. elegans} Hox genes positively [15], the \textit{C. elegans} PcG gene homologues \textit{mes}-2 and \textit{mes}-6 were not initially reported to affect cell fates regulated by Hox genes; instead, these genes were found to have a novel function in regulating X chromosome silencing in the germ line (with their mutation resulting in a \textit{male-tail} phenotype) [16]. But recent studies have indeed implicated PcG proteins as global regulators of Hox expression in \textit{C. elegans} [17\textsuperscript{**,}18\textsuperscript{**,}19].

Ross and Zarkower [17\textsuperscript{**}] performed a screen for suppressors of mutations in \textit{mab-3} (a \textit{doublesex}-related sexual regulator with roles in male tail ray specification). They found that mutations in \textit{mes}-2 and \textit{mes}-6 could suppress \textit{mab-3} phenotypes and resulted in ectopic Hox gene expression. Ross and Zarkower [17\textsuperscript{**}] propose that loss of negative regulation of Hox genes by the \textit{mes}-2 and \textit{mes}-6 PcG homologues rescues the loss of positive regulation by \textit{mab-3}. In agreement with this hypothesis, somatic expression of MES-2 and MES-6 protein was found to be widespread.

The \textit{C. elegans} homologue of the \textit{para-Hox} gene \textit{caudal}, \textit{pal-1}, is required to regulate the expression of the Hox gene \textit{mab-5} in male-tail rays. Zhang et al. [18\textsuperscript{**}] looked for \textit{pal-1} suppressors and identified \textit{suppressor of \textit{pal} (sop)} loci, including \textit{sop-2}. The \textit{sop-2} mutation on a wild-type background resulted in the ectopic expression of Hox genes, whereby Hox genes switch on at the correct time and place but are subsequently re-expressed ectopically in diverse body regions. SOP-2 contains a sterile \textit{z} motif (SAM) domain, or pointed domain, which is also present in PcG proteins and is expressed from an early stage in nearly all cells. These data suggest that \textit{sop-2} may be part of a global Hox repression system in \textit{C. elegans} [18\textsuperscript{**}]. One interpretation of this is that global Hox suppression mechanisms are conserved despite the degenerate \textit{C. elegans} cluster and deterministic development. However, SOP-2 is not an orthologue of any of the PcG genes in other animals; indeed, its SAM domain is more like those of other transcription factors.

Another gene found in Ross and Zarkower’s \textit{mab-3} suppressor screen [17\textsuperscript{**}], \textit{mes}-3, does not have any known homologues and is a novel factor in Hox regulation. Thus, significant changes in this ‘conserved’ system have evolved, including the recruitment of new elements. Finally, it is important to note that PcG control is probably redundant to other regulatory mechanisms because \textit{mes}-2, \textit{mes}-3 and \textit{mes}-6 mutants have only mild, non-embryonic, Hox-related phenotypes in specific genetic backgrounds (i.e. \textit{mab-3} and \textit{pal-1} mutants); by contrast, PcG mutations are severe in flies and vertebrates. Thus, this partially conserved global regulation has been sidelined in favour of the direct effects of other signalling systems in nematodes.

**Comparative evolution of nematode Hox gene function**

Nematodes are especially useful for evolutionary developmental biology because their conserved simple body plan allows homology to be inferred at the cellular level [20]. Nematode ‘satellite’ models have begun to shed light on macro-evolutionary developmental mechanisms in nematodes, particularly those involving the Hox genes \textit{lin-39} and \textit{mab-5} [21,22].

The \textit{lin-39} gene is required for vulval development in all species in which it has been studied [22,23]. But in \textit{Pristionchus pacificus}, a powerful genetic satellite model developed by Sommer and co-workers, significant evolutionary changes have occurred with respect to how \textit{lin-39} carries out its function. In \textit{P. pacificus}, \textit{lin-39} defines the vulval equivalence group, and vulval precursor cells (VPCs) cells lacking \textit{lin-39} activity undergo apoptosis (rather than cell fusion, as in \textit{C. elegans}). If programmed cell death is defective (such as in a \textit{lin-39} \textit{ced-3} double mutant), a normal vulva is still formed. Thus, unlike in \textit{C. elegans}, \textit{lin-39} is not required for vulval morphogenesis in \textit{P. pacificus} and instead has a role in preventing apoptosis [24].

A putative phosphorylation site and a mitogen-activated protein kinase (MAPK) docking site motif in the carboxyl terminus of \textit{C. elegans} LIN-39 were considered to be possible targets of the epidermal growth factor (lin-3), Ras (\textit{let-60}) and MAPK (\textit{mpk-1}) pathway known to regulate \textit{lin-39} specification of VPCs in \textit{C. elegans} [25], but these sequence features are absent in \textit{LIN-39} of \textit{P. pacificus}. This suggests that \textit{LIN-39} of \textit{P. pacificus} may be regulated differently and that it will not be able to substitute for the \textit{C. elegans} protein. However, \textit{P. pacificus} LIN-39 can substitute for the \textit{C. elegans} protein under the control of the \textit{C. elegans} regulatory regions [22]. Grandien and Sommer [22] also showed that the putative MAPK target motifs were not required for the function of \textit{C. elegans} LIN-39. As the \textit{P. pacificus} protein can substitute for the \textit{C. elegans} protein, it is likely that changes both in the regulatory regions and in other genes involved in vulva development are responsible for the changing role of \textit{lin-39}.

\textit{P. pacificus} and \textit{C. elegans} mutants of \textit{mab-5} are very similar with respect to their male-tail phenotype, with homologous rays (R1–R6) failing to develop in each species. Furthermore, wild-type \textit{P. pacificus} \textit{mab-5} can substitute for the \textit{C. elegans} gene when driven by regulatory regions
from *C. elegans*, restoring rays to the male tail. But *P. pacificus* mutants of *mab*-5 also have severe defects in the embryonic M lineage that result in the induction of an ectopic vulval phenotype in a posterior VPC [26]. By contrast, *C. elegans* *mab*-5 mutants do not have a vulval defect phenotype and have less severe M lineage defects. In this instance, the *mab*-5 gene seems to have evolved with respect to its role in patterning blast cell fates in the midbody, but to have conserved its role in patterning a particular structure — the male tail [21].

Conclusions: Hox genes and body patterns

The nematodes offer a new view of Hox gene functioning in body pattern evolution. Rather than being the bedrock upon which an overall plan is built, Hox genes have become ordinary bit players, retaining some (presumably ancestral) core roles, but being dismissed from others and recruited to new ones as required. In particular, the essential links among chromosomal placement, overall regulation and body part patterning have been broken. Why have nematodes achieved this, when other equally successful phyla (such as Arthropoda) are still addicted to the strict lineage-driven mechanism is an invention of one mental pattern of non-rhabditid nematodes suggest that a collinear Hox cluster? Emerging data on the evolution of lineal developmental mechanisms, and the evolution of the nematode cluster *P. pacificus* do not argue against a collinear Hox cluster? Emerging data on the evolution of lineal developmental mechanisms, and the evolution of the nematode cluster *P. pacificus* do not argue against the Hox genes from the earthworm *Perionyx excavatus*. Dev Genes Evol 2003; 213:207-210.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

• of special interest

•• of outstanding interest


This paper describes cross-species transgenesis experiments using promoter elements which demonstrate that both fly and nematode Hox genes are regulated by wg-related pathways.


Demonstration that the emerging functional genomics tool, RNA-mediated interference, can be used to modulate gene function in a major human parasite, opening the door to gene function studies across the nematodes.