*seminars in* **CELL & DEVELOPMENTAL BIOLOGY,** Vol 7, 1996: pp 539–551

# Diversification of arthropod Hox genes as a paradigm for the evolution of gene functions

*Michalis Averof, Rachel Dawes* and David Ferrier**

*Hox genes are highly conserved throughout the metazoans in both sequence, chromosomal arrangement, and function. Yet within the arthropods these genes have undergone considerable diversification. We examine ancestral and novel functions of arthropod Hox genes in an attempt to understand how these functions might evolve. We suggest that functional diversification of Hox genes begins with the acquisition of multiple distinct* cis*-regulatory elements responsible for different aspects of their gene expression. Gene duplication may serve to dissociate the functional-selective constraints associated with each of these regulatory elements and to allow divergence of the corresponding coding sequences.*

**Key words:** arthropods / evolution / gene duplication / Hox genes / regulatory elements



THE HOX GENE family is characterized by sequence motifs in and around the homeobox, clustered chromosomal organization, and a function in positional specification along the body axis. These features, which reflect the common origin and conservative nature of the Hox genes, have been extensively reviewed elsewhere.[1,2] In this review we examine the diversity of Hox gene functions in the arthropods, and focus on how these genes may acquire novel functions during evolution. As key developmental regulators Hox genes are under rigorous selective constraints. Yet their function as promiscuous transcription factors, with loose target recognition and multiple interacting cofactors, offers wide-ranging possibilities for the evolution of new functions. Like the globins (an established model for the study of gene duplication and diversification), they comprise a family of genes where the history of gene duplications can be traced by comparing the sequences and chromosomal arrangement of existing genes.

*From EMBL, Meyerhofstrasse 1, Heidelberg 69117, Germany and *Wellcome/CRC Institute, Tennis Court Road, Cambridge, CB2 1QR, UK*



*1084-9521/96/040539 + 13 $18.00/0*

### *The 'canonical' function of arthropod Hox genes*

In all metazoans that have been investigated to date (including nematodes, annelids, arthropods and chordates) Hox genes are thought to play a role in defining the positional identity of structures along the anteroposterior body axis.[3] Genetic and developmental evidence suggest that in the arthropods, Hox genes operate within the segmental framework of the body to confer regional identity to most segmental structures[4,5] (Table 1). This identity is directly related to the regionalized expression of these genes.

### *Evolutionary patterns and the origins of arthropod Hox genes*

Comparison of Hox genes between disparate animal groups indicates that, although the Hox family as a whole appears to be conserved, individual members of the family need not be. Gene duplication and diversification and/or gene loss appear to have changed the structure of Hox clusters during the evolution of different animal groups.[6] By comparing the insect and vertebrate Hox clusters we can, somewhat arbitrarily, divide Hox genes into four groups which exemplify four different evolutionary patterns (summarized in Figure 1).†

*Anterior-acting 'head' genes (*lab/Hox1, pb/Hox2, Hox3, Dfd/Hox4*) — conservation*

Each of these distinct gene classes shows high conservation of gene structure and function (with the

*†Hox clusters have undergone some large-scale structural changes during evolution. In some higher insects the Hox cluster (HOM-C) has been split in two parts, BX-C and ANT-C. Within the vertebrate lineage, large-scale duplications have given rise to four clusters with largely overlapping sets of genes.[1,7] Corresponding members of each cluster show extensive conservation in sequence, expression and function, constituting sets of genes which are termed paralogous groups. In this review, when referring to vertebrate Hox genes we refer to entire paralogous groups.*

**Table 1.** Comparative data on canonical functions of arthropod Hox genes

| Gene name | Arthropod group | | Available data | | Inferred homeotic function | Refs |
|---|---|---|---|---|---|---|
| | | | expr. | genetic | | |
| lab | Insect | Diptera | + | + | Ic | 5, 57, 70 |
| pb | Insect | Diptera | + | + | Mx1, Mx2 | 5, 56, 70 |
| | | Coleoptera | – | + | Mx1, Mx2 | 71, 72 |
| Dfd | Insect | Diptera | + | + | Md, Mx1 | 5, 70 |
| | | Hymenoptera | + | – | Md, Mx1 | 73, 74 |
| Scr | Insect | Diptera | + | + | Mx2, T1 | 5, 70 |
| | | Coleoptera | – | + | Mx2, T1 | 71, 72 |
| | | Orthoptera | + | – | Mx2, T1 | 75 |
| Antp | Insect | Diptera | + | + | T1p–T3a | 5, 55 |
| | | Coleoptera | – | + | T1p–T3a | 71, 72 |
| | | Orthoptera | + | – | T1–T3 | 76 |
| | Crustacean | Anostraca | + | – | Mx2, thorax | 35 |
| Ubx | Insect | Diptera | + | + | T2p–A8 | 4, 77 |
| | | Lepidoptera | – | + | T2p–A10? | 78 |
| | | Coleoptera | – | + | T2p–A10? | 71, 72 |
| | | Orthoptera | + | – | T2p–A10? | 79 |
| | Crustacean | Anostraca | + | – | Thorax | 35 |
| abdA | Insect | Diptera | + | + | A1p–A8 | 4, 36 |
| | | Lepidoptera | + | + | A1p–A10 | 15, 78, 80 |
| | | Coleoptera | + | + | A1p–A8, –A10? | 71, 72, 81 |
| | | Orthoptera | + | – | A1p–A10 | 82 |
| | Crustacean | Anostraca | + | – | Thorax | 35 |
| AbdB | Insect | Diptera | + | + | A4p–A8a, A8p–A9 | 4, 16 |
| | | Coleoptera | – | + | Posterior abdomen? | 72 |
| | | Orthoptera | + | – | A8p–A11 | 19 |
| | Crustacean | Anostraca | + | – | Genital | 35 |

Available genetic and expression data are presented for different arthropod groups (mainly insects). The inferred function of each gene is indicated in terms of the segmental identity specified (Ic=intercalary, Md=mandibular, Mx1=first maxillar, Mx2=second maxillar or labial, Tn=n$^{th}$ thoracic, An=n$^{th}$ abdominal). Due to space limitations only selected references could be given for *Drosophila* genes

exception of the insect Hox3, see later). Their sequence conservation and relative order along the chromosome allows us to assign homology unambiguously between individual insect and vertebrate 'head' genes.[8] This structural conservation is matched by functional conservation in patterning the head region, although the divergence of insect and vertebrate morphology does not allow a more detailed comparison of their gene functions in the head.

*Middle-acting 'trunk' genes (*Scr, Antp, Ubx, abdA *and* Hox5,6,7,8*) — independent diversification*

Sequence comparisons identify these 'trunk' genes as a closely related set of Antp-like genes (also closely related to *ftz*, discussed below, and the 'head' genes *Dfd* / Hox4[6]). However, sequence comparisons do not allow unambiguous assignment of a one-to-one homology between individual vertebrate and insect 'trunk' genes.[6] A minimal hypothesis would therefore suggest that these genes originated from a single gene precursor in the last common ancestor of vertebrates and insects. While it remains possible that more than one 'trunk' gene existed in the ancestral cluster (see Figure 1), it is clear that these genes have acquired most of their specific sequence characteristics independently, in the lineages leading to vertebrates and insects. It has been suggested that this diversification was associated with the evolution of segmental diversity in the trunk region of the insects.[4,9]

*Posterior-acting 'tail' genes (*AbdB *and* Hox9-13*) — gene duplication or loss*

On the basis of their sequence, these genes form a distinct and divergent class of Hox genes.[8] A single gene of this class has been found in several arthropods studied to date (Table 1), whereas at least five distinct paralagous groups are found in vertebrates.[8] It is usually assumed that the last common ancestor of vertebrates and insects possessed a single *Abd-B* gene, which subsequently duplicated in the lineage leading to the vertebrates.[6] However, it is also plausible that several *AbdB* genes were present in that common ancestor and lost in the lineage leading to the insects.

*Divergent Hox genes (*zen, bcd, ftz*) — divergence (see Table 2A)*

Several homeobox-containing genes interspersed

within the arthropod Hox clusters have diverged so rapidly that their relationship to other Hox genes is obscured. These genes appear to have acquired radically new functions which are restricted to the arthropods, and sometimes even to specific insect groups (discussed later).

In addition to these genes are several others which, although not presently associated with the insect Hox clusters, are thought to derive from the same ancestral type of homeobox gene. Their relationship to Hox genes is inferred on the basis of their sequence and conserved roles in specification of anterior and posterior regions of the body in vertebrates, insects, and possibly nematodes. These genes include *otd, ems, cad* and *eve* (see Table 2B). Intriguingly, *eve* and *ems* homologues are located near the Hox clusters in vertebrates and nematodes respectively.[10,11] A number of additional genes which are not Hox genes (as they are not located in the Hox clusters) but whose sequences place them within the Hox family, have been identified in vertebrates.[8] These include *gsh1*[12] and *Xlhbox8*.[13] Homologues of these genes are likely to be present in diverse animal groups[14] but have not so far been identified in arthropods.

Such comparisons between divergent animal groups indicate that the last common ancestor of all arthropods already possessed an extensive complement of Hox genes: the 'head' genes, at least one 'trunk' gene and at least one 'tail' gene. Comparisons between different arthropod subgroups identify how this ancestral set of Hox genes has changed during the subsequent evolution of the arthropods.

## Hox gene function and evolution in arthropods

### *Variations of the 'canonical' function*

*Comparisons within insects*

The conserved body plan of the insects is paralleled by conservation of homeotic gene number, sequence, and function in insect species studied to date (see Table 1). Their domains of gene expression have also been broadly conserved, although some interesting variations in Hox gene regulation have been noted. We consider briefly two examples.

The first concerns the development of abdominal prolegs, rudimentary appendages that develop in abdominal segments of some insect larvae. *AbdA* and

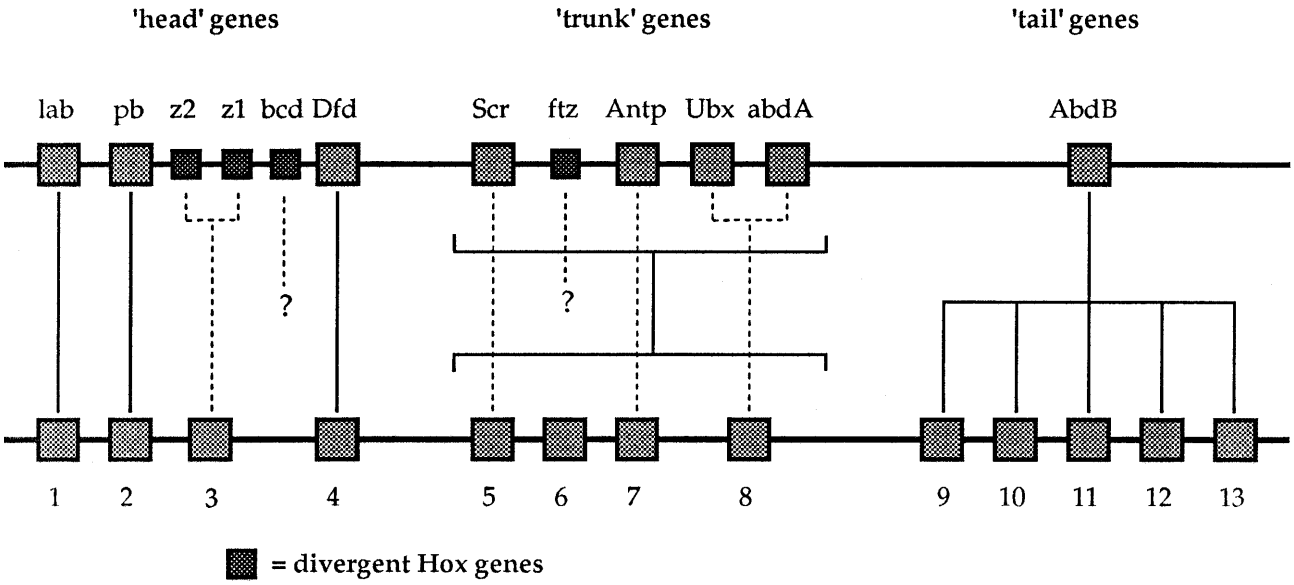


**Figure 1.** Patterns of Hox gene diversification. Comparisons between insect (upper) and vertebrate (lower) Hox clusters identify four patterns of Hox gene diversification, as described in the text: conservation of 'head' genes; independent diversification of 'trunk' genes; duplication or loss of 'tail' genes and divergence of non-homeotic Hox genes. Putative relationships between individual insect and vertebrate 'trunk' genes (based on limited sequence similarities and some data on functional conservation[83]) are indicated by dashed lines. Vertebrate genes are named by paralagous group number and fly genes are named by gene symbol (z1 and 2 indicate *zen1* and *2*).

**Table 2.** Comparative data on (A) divergent Hox genes, and (B) non-Hox genes with Hox gene characteristics

| Gene name | Animal group | | Available data: expr. | Available data: genetic | Hexapeptide | In Hox cluster | Inferred function | Refs |
|---|---|---|---|---|---|---|---|---|
| (A) | | | | | | | | |
| zen-Hox3 | Insect | Diptera | + | + | – | + | Amnioserosa, optic lobe | 5, 42, 44 |
| | | Coleoptera | + | – | – | + | Amnion, serosa | 43 |
| | | Orthoptera | + | – | – | + | Amnion, serosa | 43 |
| | Vertebrate | | + | + | + | + | Homeotic | 84, 85 |
| bcd | Insect | Diptera | + | + | – | + | Early anterior | 47–49,86–88 |
| ftz | Insect | Diptera | + | + | – | + | Segment, CNS | 40,41,89–90 |
| | | Coleoptera | + | – | + | + | Segment? CNS, post. | 54 |
| | | Orthoptera | + | – | + | + | CNS, posterior | 53 |
| (B) | | | | | | | | |
| eve | Insect | Diptera | + | + | – | – | Segmentation, CNS | 91–94 |
| | | Coleoptera | + | – | ? | ? | Segment., CNS, post. | 54, 95 |
| | | Grasshopper | + | – | – | ? | CNS, posterior | 96 |
| | Vertebrate | | + | + | – | + | CNS, posterior | 10, 97 |
| | Nematode | | + | + | ? | – | Posterior | 98 |
| cad | Insect | Diptera | + | + | + | – | Posterior, gut | 99, 100 |
| | | Lepidoptera | + | – | + | ? | Posterior, gut | 101 |
| | Vertebrate | | + | + | + | – | Posterior, gut | 102–105 |
| | Nematode | | – | + | + | – | Posterior | 8, 106 |
| ems | Insect | Diptera | + | + | + | – | Head | 107, 108 |
| | Vertebrate | | + | – | + | – | Head | 109, 110 |
| otd | Insect | Diptera | + | + | – | – | Head | 108, 111 |
| | Vertebrate | | + | + | – | – | Head | 109,110,112 |

Available genetic and expression data are presented for different animal groups. The inferred function of each gene is indicated in processes like segmentation, neurogenesis (CNS), and regional specification in various parts of the body (brain, head, posterior). The presence or absence of a conserved 'hexapeptide motif'[8] and known linkage to the Hox clusters are also noted. Due to space limitations only selected references could be given for some genes

*Ubx* are co-expressed in the abdomen of insects, where they usually repress the development of legs. It has recently been shown that in butterflies *abdA* and *Ubx* become down-regulated in specific parts of abdominal segments, presumably to allow development of prolegs.[15]

The second example concerns changes in *AbdB* expression. In *Drosophila* the *AbdB* gene is active in two domains of the abdomen (A4p–A8a and A8a-backwards), each under the control of distinct *cis*-regulatory elements.[16-18] Expression in the most posterior domain is conserved in locusts, but the more anterior expression appears to be absent. It seems likely that the regulatory elements associated with the anterior function were acquired by *Abd-B* during the evolution of specific insect groups and may be associated with subtle changes in abdominal morphology.[19]

Further understanding of homeotic gene function will surely lead to new questions, to be addressed by comparison of closely related insect species. Recent studies in *Drosophila*, for example, highlight the importance of precise spatial and temporal regulation of Hox genes within their domains of gene expression.[20,21] Regulatory changes at such a fine level may indeed be relevant to morphological variations within the insect lineage.

*Comparisons between crustaceans and insects*

In spite of past controversy, crustaceans are now thought to be close relatives of insects,[22-24] and segmentation is almost certainly homologous in these two groups.[25] Apart from a shared organization of head segments, however, insect and crustacean body plans are quite divergent. Various regions of the trunk appear to have diversified independently during the evolution of insects and different crustacean subgroups; terms like 'thorax' and 'abdomen' are used in a descriptive sense, rather than to designate homologous parts.[26] Thus, the study of 'trunk' Hox genes is interesting, to determine the extent to which segment diversification in the arthropod trunk has been associated with the evolution of Hox genes.

The isolation of distinct and unambiguously-identified homologues of all known insect 'trunk' genes (*Scr*, *Antp*, *Ubx*, *abdA*) from a branchiopod crustacean

(*Artemia franciscana*)[27] suggests that 'trunk' gene duplications pre-date segment specialization within the insect and crustacean trunk. Preliminary data from a horseshoe crab (*Limulus polyphemus*)[28] indicate that distinct 'trunk' genes probably evolved even before the divergence of chelicerates, crustaceans and insects. Moreover, extensive work on leech Hox (*Lox*) genes has shown that at least one of the 'trunk' gene duplications, between *Antp*/*Scr*-like (leech homologues *Lox1, Lox5*) and *Ubx*/*abdA*-like (leech homologues *Lox2, Lox4*) genes, occurred before the divergence of the annelids and arthropods.[29-32] Limited data from other annelids, a polychaete (*Ctenodrilus serratus*)[33] and an oligochaete (*Stylaria lacustris*),[34] are consistent with these observations. From these data it is clear that the gene duplications which gave rise to the insect 'trunk' genes occurred long before, and cannot be directly associated with the diversification of segments in the insect trunk.

Knowing that an extensive complement of Hox genes is shared between crustaceans and insects raises questions about the function of these genes in the ancestors and relatives of the insects: Were they always involved in specifying distinct segmental types, and how did their function change during segment diversification in the insect trunk? The most relevant comparative data come from *Artemia*.[35] In *Artemia*, Hox genes *Antp, Ubx* and *abdA* are co-expressed in the thoracic region, unlike in insects where these genes are expressed in distinct regional domains within the trunk. Since there are no obvious differences in segmental morphology within the thorax of *Artemia* (the thorax is a 'homonomous' region of the body), these genes may collectively specify a common thoracic segmental identity. This implies that related Hox genes, like *Antp, Ubx* and *abdA*, may not always be used to specify distinct segmental identities.

If the 'canonical' function of specifying thorax or trunk is shared between the *Antp, Ubx* and *abdA* genes in *Artemia*, we must ask whether the functions of these genes are likely to be truly redundant. It is unclear whether the products of these genes have acquired different properties in *Artemia*. Preliminary observations, however, suggest that they are expressed in somewhat different subsets of cells within the thorax (for example, unlike *Antp, abdA* is predominantly expressed within the nervous system). Thus these genes may have acquired distinct functions due to differential intra-segmental regulation, even if their coding sequences retain similar properties. A similar conclusion can be reached for aspects of *abdA* and *Ubx* function in the *Drosophila* abdomen,[20,36] and for paralogous groups of Hox genes in the vertebrate hindbrain.[37-39]

How then, do we explain the evolution of distinct segment types within the insect trunk? The differences in gene expression observed between the insect and *Artemia* 'trunk' genes suggests that this event may be associated with the resolution of expression patterns of 'trunk' genes into distinct regional domains, and with changes in the way these genes regulate their downstream targets.[35]

### *Divergent Hox genes with novel functions*

A number of Hox genes appear to have lost their ancestral homeotic function and acquired entirely new functions within the arthropod lineage. In this section we present three examples. We examine the origin of these 'runaway' Hox genes and speculate on how such novel functions might arise.

*Specific examples*

*Fushi tarazu (ftz)* provides the best studied example of a divergent Hox gene within the arthropod lineage. Two major functions have been identified for *ftz* in *Drosophila*, one in establishing the pattern of segments in the early embryo (a pair-rule function) and a later one in specifying neuronal types within the central nervous system (CNS).[40,41] The gene is located in the Hox cluster between *Antp* and *Scr*, and its sequence is divergent, although clearly related to the Antp class of 'trunk' genes.[8]

*Ftz* homologues have been identified in flies, a beetle and a locust (Table 2A). These homologues share certain loosely conserved sequence motifs and a strikingly similar pattern of expression in the CNS, which suggest that a distinct *ftz* gene with the characteristic neural expression must have arisen before the divergence of major insect groups. The early expression and associated function in segmentation, however, appear to have changed significantly during insect evolution: a modified pair-rule expression is still observable in the beetle but absent in the locust. In both species early expression is restricted to a posterior portion of the embryonic primordium which corresponds roughly to the prospective trunk. Intriguingly, the early and late expression patterns of *ftz*, which show different degrees of conservation, are controlled by separable regulatory elements in *Drosophila* (see Figure 2).

Putative *ftz* homologues have also been identified in a crustacean (*AfHx1*)[27] and a leech (*Lox1*).[30] The *Lox1* gene shows closest sequence similarity to Hox6, *ftz*

and *Scr*, and its domain and pattern of expression in the CNS shows some similarities to *ftz*. These observations suggest that *ftz* and *Lox1* may derive from an ancestral *Scr*- or Hox6-like 'trunk' gene which already had a specific function in the CNS before the arthropod-annelid divergence.

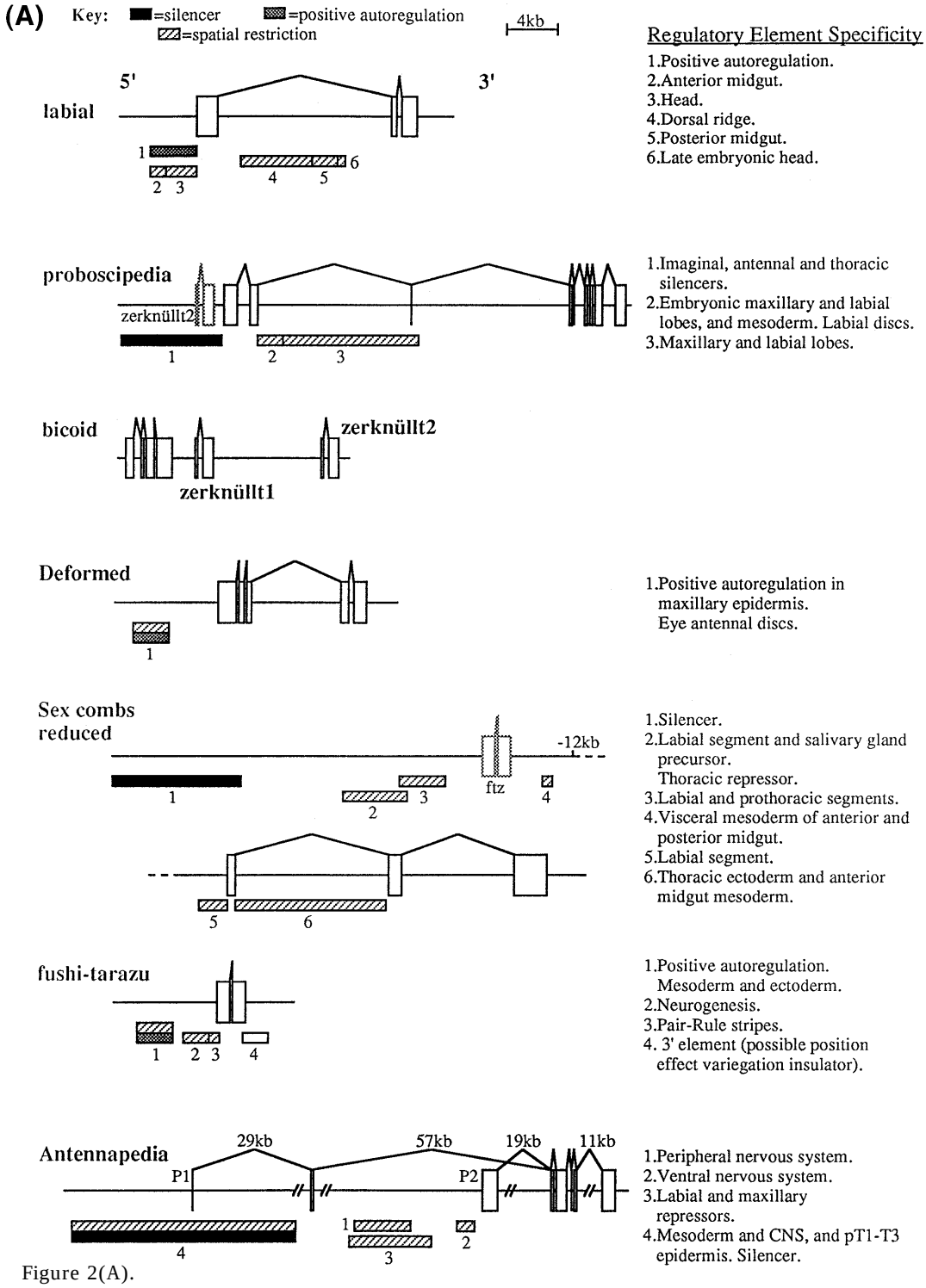


Figure 2(A).

*Zerknüllt (zen)* is a locus involved in specifying dorsally-derived structures of the early *Drosophila* embryo (extra-embryonic membranes and optic lobes).[42] Its chromosomal location, between *pb* and *Dfd*, and sequence provide some hint of a relationship with vertebrate Hox3.

Homologues of *Drosophila zen* have recently been identified in a locust and a beetle (ref 43, Table 2A); these genes share some sequence features with *Drosophila zen* and, like *zen*, are expressed in extra-embryonic membranes. Interestingly, their homeodomains show striking sequence similarity to vertebrate Hox3 genes. This observation suggests that *zen* genes derive from an ancestral Hox3 gene. It remains unclear when and how *zen* genes diverged from more canonical Hox3 genes, acquiring new functions in the extra-embyronic membranes. A putative *zen*-Hox3 homologue, with unknown function, has also been identified in a chelicerate.[28]

The *Drosophila melanogaster zen* locus comprises two adjacent and closely related genes, *zen1* and *zen2*. The two genes are expressed in a very similar pattern, but all known functions of the *zen* locus are effectively mediated by *zen1*.[44] Sequence comparisons and the absence of *zen2* from closely related species (*D. pseudobscura, D. subobscura*) suggest that *zen2* has arisen by a very recent gene duplication and has diverged rapidly ever since.[45,46] Although no known function has been associated with *zen2*, the integrity of its coding sequence and homeobox motif, despite rapid sequence divergence, are consistent with it being functional.

*Bicoid (bcd)* is involved in patterning the anterior of the early *Drosophila* embryo.[47,48] The gene is located in the Hox cluster between *Dfd* and the *zen* genes, but its sequence and function have diverged so extensively from those of other Hox genes that its origin and relationships remain unclear.[8] We can only speculate that it derives from duplication of adjacent Hox genes followed by rapid diversification. Divergent *bcd* homologues have been found in other Diptera but not in any other insects (Table 2A), and functional assays

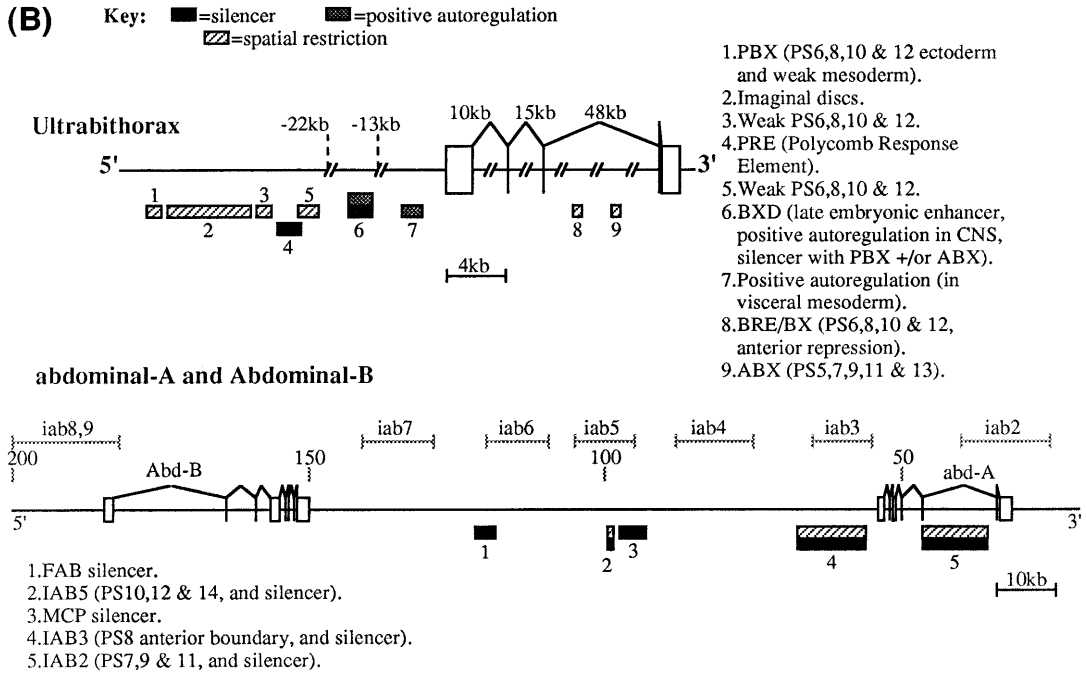


**Figure 2.** The Hox genes of *D. melanogaster*, indicating the position and modular nature of regulatory elements, as characterized by lacZ reporter constructs. Open boxes represent exons. The shaded boxes beneath the genes represent fragments with the specified activities in lacZ reporter constructs. (A) Regulatory elements of the Antennapedia complex (ANT-C). (B) Regulatory elements of the Bithorax complex (BX-C). The iab regions depicted above the *abd-A* and *Abd-B* genes are genetically defined regulatory regions, producing parasegment-specific expression. Note that the scale of *abd-A*/*Abd-B* diagram is different from the scale of the others. (References; *labial*,[113] *proboscipedia*,[114] *Deformed*,[115] *Sex combs reduced*,[58] *fushi tarazu*,[62,116] *Antennapedia*,[117,118] *Ultrabithorax*,[119-124] *abdominal-A*,[60,125] and *Abdominal-B*,[126,127]).

have failed to reveal any *bcd*-like activity in the anterior cytoplasm of non-dipteran and even divergent dipteran species.[49] There is therefore no evidence that *bcd* function, as has been described in *Drosophila*, exists at all outside the Diptera.

*Relating canonical and divergent functions of Hox genes*

If new Hox gene functions have evolved without major disruption to developmental processes they must have been acquired in parallel and in addition to ancestral homeotic functions. A possible exception, the origin of *zen* from an ancestral Hox3 gene (which it appears to have totally replaced in insects), represents an interesting puzzle.

The evolution of divergent functions has involved changes at the level of gene regulation, interactions with putative co-factors, and specificity for downstream promoters. This is evident from differences in gene expression, altered constraints on protein sequences, differences in target specificities, etc. (Table 2 and refs 50-52). Despite such differences in 'runaway' Hox genes, however, some aspects of their function and expression may be directly related to those of canonical Hox genes. The early function of *ftz* in locust and beetle, for example, showing a sharp anterior boundary of expression,[53,54] may have been inherited from an ancestral gene with homeotic functions. Moreover, there is evidence to suggest that some non-homeotic functions are carried out by *bona fide* homeotic genes. For example, the expression of *Antp, lab* and *pb* in specific cells within the CNS, throughout the body axis,[55-57] is strongly reminiscent of the *ftz* expression within the CNS. These expression patterns are often controlled by separable enhancers (See Figure 2), which could therefore be recruited for the regulation of new genes (*Drosophila ftz* is actually embedded within the regulatory elements of *Scr*[58]). Another example is the function of *lab* in specifying a particular cell type within a particular region of the gut.[59]

The canonical function of Hox genes has been defined as the specification of regional identity, rather than specification of particular structures or cell types; yet it is certainly the case that Hox genes participate in developmental pathways which ultimately lead to decisions on cell fate. A better understanding of Hox gene functions at the cellular level may clarify the extent to which 'canonical' and these so-called 'divergent' functions are related.

### *Chromosomal arrangement and gene regulation*

The insect Hox clusters are very large: 650 kb in *Drosophila*, and a similar size in *Bombyx* and *Schistocerca*[5,60] (Y. Suzuki pers. comm.; D. Ferrier, in prep.). Coding sequences occupy only a small fraction of that size, and the extensive stretches of non-coding DNA contain sequence elements important for gene regulation.

Figure 2 summarizes our current knowledge on the organisation of regulatory elements of the *Drosophila* Hox genes. These elements (like those of other genes) can be dissociated into independent modules controlling different aspects of the gene's expression: parasegmental expression, autoregulation, silencing, and germ layer and tissue specificity. The individual modules appear to be relatively flexible, since they produce the same expression patterns in various genomic contexts and with different promoters.[18,61,62] Moreover, they can be associated combinatorially to regulate complex patterns of expression from individual promoters. Thus, the modular nature and flexibility of these elements provide extensive opportunity for regulatory evolution, particularly since this allows different aspects of gene regulation to evolve independently.

In the *Drosophila* BX-C genes, expression in individual abdominal parasegments is driven by separate regulatory elements which behave genetically as separate genes (this, in fact, was the original suggestion made by Lewis[4]). Acquisition of such independent and dissociable elements provides a way of increasing functional diversity at the level of gene regulation, a mechanism which closely parallels the evolution and diversification of entire genes. This parallel can be illustrated by the evolution of *Abd-B* genes: In chordates functional diversity of the *Abd-B* genes is reflected by increasing gene number (Figure 1), whereas in *Drosophila* functional diversity has been achieved by the acquisition of independent region-specific regulatory elements associated with a single *Abd-B* coding sequence (Figure 2).[63]

## The evolution of gene functions: the role of modular regulatory elements and gene duplication

In this section we consider the nature of gene duplications, the selective and mutational forces that determine their early fate, and the immediate and long-term consequences for the evolution of gene functions. Our discussion is based primarily on thoughts about the evolution of arthropod Hox

genes, but should be relevant to the evolution of any gene family over long evolutionary timescales.

Our idealized gene consists of a coding sequence with its associated regulatory elements. We envisage this coding sequence as a discrete and indivisible unit giving rise to a specific gene product (for the purposes of this discussion we temporarily ignore alternate isoforms and the multi-domain constitution of some proteins). In contrast, we expect that multiple regulatory elements with diverse and dissociable activities (tissue- or cell-specific, temporally-restricted, silencing or activating) are interspersed over long distances in and around this coding sequence. This resembles the organization of Hox genes (Figure 2) and other genes whose regulation has been studied in some detail (e.g. beta-1 tubulin and string (cdc25) in *Drosophila*[64,65]).

Processes like unequal recombination can duplicate large contiguous pieces of DNA. The nature of such processes suggest that the boundaries of duplicated regions should be arbitrary with respect to underlying regulatory and coding information, and that the duplicated fragments will be accurate copies of each other. Some of these duplications will therefore include perfect copies of entire genes (including coding sequence and all the associated regulatory elements), whereas others will include accurate but partial copies, possibly missing some of the regulatory elements associated with the gene (Figure 3). In either case, it is unlikely that the gene duplication event itself would qualitatively affect the coding sequence; a pair of newly-duplicated genes should encode identical amino acid sequences, with identical biochemical properties.

The immediate evolutionary fate of duplicated genes will depend on the relative forces of selection and mutation. In the absence of selective pressure to maintain both copies (if their functions remain truly redundant), one will inevitably accumulate deleterious mutations and degenerate into a pseudogene. In some cases a selective advantage associated with increased gene dosage may be sufficient to maintain two copies of the gene; we consider this to be unlikely for genes whose products are not required in bulk (unlike so-called 'housekeeping' or structural proteins or RNAs). There are two additional ways in which newly-duplicated copies of a gene can be individually selected and maintained; these depend on the acquisition of distinct, non-overlapping functions. The first involves diversification of their coding sequences, and the second involves differences in gene expression. The first possibility seems highly improbable if we consider that random mutations are far more likely to be deleterious than to confer a novel function. The second appears plausible, since we already know that the gene duplication event itself can lead to differential partitioning of regulatory elements (Figure 3). Thus, we expect those newly-duplicated genes which are selectively maintained to have almost identical gene products expressed in at least partially non-overlapping patterns.

In the long-term these genes should now be free to evolve as independent units. If the functions of both copies happen to be under the same selective pressures as their progenitor we expect their sequences to

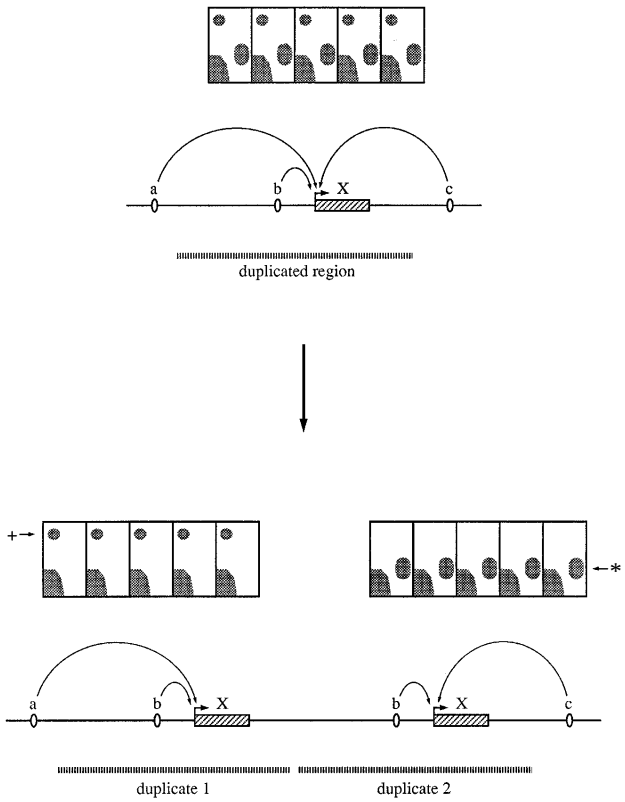


**Figure 3.** Gene duplication dissociates regulatory elements and functions. In our example, gene *X* is expressed in a complex metameric pattern controlled by several distinct regulatory elements (*a*, *b*, *c*, etc.). A tandem duplication creates two identical coding sequences, each of which comes under the influence of different subsets of nearby regulatory elements. The overall expression and function of gene *X* remains unchanged, but each of the duplicate coding sequences acquires unique non-overlapping functions by virtue of its expression in different populations of cells (marked + and *), for which they are selectively maintained. Dissociation of these functions into different genes allows their subsequent independent evolution.

be constrained in the same manner and to evolve by slow accumulation of mostly neutral mutations. The products of this process may be exemplified by pairs or small groups of genes whose sequence divergence appears not to be functionally important and whose distinct functions are defined primarily by differential expression (e.g. mouse *HoxA4, B4* and *D4*,[37] *En-1* and *En-2*;[66] *C. elegans APX-1* and *LAG-2*;[67] *Drosophila prd, gsb* and *gsbn*,[68] *kni* and *knr*.[69]) In other cases, however, gene duplication may dissociate functions that were previously constrained by having to act through a single coding sequence, and thus create new opportunities for change. In these cases, gene duplication may be followed by rapid sequence divergence and the evolution of new functions (as described for *ftz, zen* and *bcd*).

## Conclusions

We have discussed possible routes and constraints for the evolution of new functions on the basis of comparative data from arthropod Hox genes. We have suggested that gene duplications are not directly associated with the evolution of new genetic functions and therefore need not be associated with discontinous jumps in morphology ('hopeful monsters'). We believe, however, that gene duplications create a potential for the evolution of gene functions by dissociating different functional constraints. Thus, they allow functions driven by different regulatory elements to evolve independently and to diversify. We consider the modular organization of regulatory regions and dissociability of these modules to be of primary importance in allowing this dissociation of constraints.

## Acknowledgements

We are grateful to Michael Akam in whose lab these thoughts have developed. We thank Michael Akam and David Stern for comments on the manuscript. Our work has been supported by the Wellcome Trust and the MRC.

## References

1. McGinnis W, Krumlauf R (1992) Homeobox genes and axial patterning. Cell 68:283-302
2. Manak RJ, Scott MP (1994) A class act: conservation of homeodomain protein functions. Development Supplement 61-71
3. Slack JMW, Holland PWH, Graham CF (1993) The zootype and the phylotypic stage. Nature 361:490-492
4. Lewis EB (1978) A gene complex controlling segmentation in *Drosophila*. Nature 276:567-570
5. Kaufman TC, Seeger MA, Olsen G (1990) Molecular and genetic organization of the Antennapedia gene complex of *Drosophila melanogaster*. Adv Genet 27:309-362
6. Schubert FR, Nieselt-Struwe K, Gruss P (1993) The Antennapedia-type homeobox genes have evolved from three precursors separated early in metazoan evolution. Proc Natl Acad Sci USA 90:143-147
7. Garcia-Fernàndez J, Holland PWH (1994) Archetypal organization of the amphioxus *Hox* gene cluster. Nature 370:563-566
8. Bürglin TR (1994) A comprehensive classification of homeobox genes, in Guidebook to the Homeobox Genes (Duboule D, ed.) pp 27-71. Oxford University Press, Oxford
9. Akam M, Dawson I, Tear G (1988) Homeotic genes and the control of segment diversity. Development Supplement 123-134
10. Dollé P, Fraulob V, Duboule D (1994) Developmental expression of the mouse *Evx-2* gene: relationship with the evolution of the HOM/Hox complex. Development Supplement 143-153
11. Wang BB, Muller-Immergluck MM, Austin J, Robinson NT, Chisholm A, Kenyon C (1993) A homeotic gene cluster patterns the anteroposterior body axis of *C. elegans*. Cell 74:29-42
12. Valerius MT, Li H, Stock JL, Weinstein M, Kaur S, Singh G, Potter SS (1995) *gsh-1* — a novel murine homeobox gene expressed in the central nervous system. Developmental Dynamics 203:337-351
13. Wright CVE, Schnegelsberg P, De Robertis EM (1988) *XlHbox8*: a novel Xenopus homeo protein restricted to a narrow band of endoderm. Development 104:787-794
14. Wysocka-Diller J, Aisemberg GO, Macagno ER (1995) A novel homeobox cluster expressed in repeated structures of the midgut. Dev Biol 171:439-447
15. Warren RW, Nagy L, Selegue J, Gates J, Carroll S (1994) Evolution of homeotic gene regulation and function in flies and butterflies. Nature 372:458-461
16. DeLorenzi M, Bienz M (1990) Expression of Abdominal-B homeoproteins in *Drosophila* embryos. Development 108:323-329
17. Casanova J, Sánchez-Herrero E, Morata G (1986) Identification and characterization of a parasegment specific regulatory element of the *Abdominal-B* gene of *Drosophila*. Cell 47:627-636
18. Celniker SE, Sharma S, Keelan DJ, Lewis EB (1990) The molecular genetics of the bithorax complex of *Drosophila*: cis-regulation in the *Abdominal-B* domain. EMBO J 9:4277-4286
19. Kelsh R, Dawson I, Akam M (1993) An analysis of Abdominal-B expression in the locust *Schistocerca gregaria*. Development 117:293-305
20. Castelli-Gair J, Greig S, Micklem G, Akam M (1994) Dissecting the temporal requirements for homeotic gene function. Development 120:1983-1995
21. Castelli-Gair J, Akam M (1995) How the Hox gene *Ultrabithorax* specifies two different segments: the significance of spatial and temporal regulation within metemeres. Development 121:2973-2982
22. Averof M, Akam M (1995) Insect-crustacean relationships: insights from comparative developmental and molecular studies. Phil Trans Roy Soc Lond (Biol) 347:293-303
23. Boore JL, Collins TM, Standon D, Daehler LL, Brown WM

(1995) Deducing the pattern of arthropod phylogeny from mitochondrial DNA rearrangements. Nature 376:163-165
24. Friedrich M, Tautz D (1995) Ribosomal DNA phylogeny of the major extant arthropod classes and the evolution of myriapods. Nature 376:165-167
25. Patel NH, Kornberg TB, Goodman CS (1989) Expression of *engrailed* during segmentation in grasshopper and crayfish. Development 107:201-212
26. Calman WT (1909) Crustacea. A Treatise on Zoology 7 (Lankester ER, ed.). Adam and Charles Black, London
27. Averof M, Akam M (1993) *HOM/Hox* genes of *Artemia*: implications for the origin of insect and crustacean body plans. Curr Biol 3:73-78
28. Cartwright P, Dick M, Buss LW (1993) *HOM/Hox* type homeoboxes in the chelicerate *Limulus polyphemus*. Mol Phyl Evol 2:185-192
29. Shankland M, Martindale MQ, Nardelli-Haefliger D, Baxter E, Price DJ (1991) Origin of segmental identity in the development of the leech nervous system. Development Supplement 2:29-38
30. Aisemberg GO, Macagno ER (1994) *Lox1*, an Antennapedia-class homeobox gene, is expressed during leech gangliogenesis in both transient and stable central neurons. Dev Biol 161:455-465
31. Wong VY, Aisemberg GO, Wen-Biao G, Macagno ER (1995) The leech homeobox gene *Lox4* may determine segmental differentiation of identified neurons. J Neurosci 15:5551-5559
32. Nardelli-Haefliger D, Shankland M (1992) *Lox2*, a putative leech segment identity gene, is expressed in the same segmental domain in different stem cell lineages. Development 116:697-710
33. Dick MH, Buss LW (1994) A PCR-based survey of homeobox genes in *Ctenodrillus serratus* (Annelida: Polychaeta). Mol Phyl Evol 3:146-158
34. Snow P, Buss LW (1994) HOM/Hox-type homeoboxes from *Stylaria lacustris* (Annelida: Oligochaeta). Mol Phyl Evol 3:360-364
35. Averof M, Akam M (1995) *Hox* genes and the diversification of insect and crustacean body plans. Nature 376:420-423
36. Macias A, Casanova J, Morata G (1990) Expression and regulation of the *abd-A* gene of *Drosophila*. Development 110:1197-1207
37. Gaunt SJ, Krumlauf R, Duboule D (1989) Mouse homeo-genes within a subfamily, Hox-1.4, -2.6 and -5.1, display similar anteroposterior domains of expression in the embryo, but show stage- and tissue-dependent differences in their regulation. Development 107:131-141
38. Graham A, Maden M, Krumlauf R (1991) The murine Hox2 genes display dynamic dorsoventral patterns of gene expression during central nervous system development. Development 112:255-264
39. Hunt P, Gulisano M, Cook M, Sham M-H, Faiella A, Wilkinson D, Boncinelli E, Krumlauf R (1991) A distinct *Hox* code for the branchial region of the vertebrate head. Nature 353:861-864
40. Wakimoto BT, Turner FR, Kaufman TC (1984) Defects in embryogenesis in mutants associated with the Antennapedia gene complex of *Drosophila melanogaster*. Dev Biol 102:147-172
41. Doe CQ, Hiromi Y, Gehring WJ, Goodman CS (1988) Expression and function of the segmentation gene *fushi tarazu* during *Drosophila* neurogenesis. Science 239:170-175
42. Rushlow C, Levine M (1990) Role of the *zerknüllt* gene in dorsal-ventral pattern formation in *Drosophila*. Adv Genet 27:277-307
43. Falciani F, Hausdorf B, Schroeder R, Akam M, Tautz D, Dennell R, Brown S (1996) Class 3 Hox genes in insects and the origin of *zen*. PNAS, in press
44. Rushlow C, Doyle H, Hoey T, Levine M (1987) Molecular characterization of the *zerknüllt* region of the Antennapedia gene complex in *Drosophila*. Genes Dev 1:1268-1279
45. Randazzo FM, Seeger MA, Huss CA, Sweeney MA, Cecil JK, Kaufman TC (1993) Structural changes in the Antennapedia complex of *Drosophila pseudoobscura*. Genetics 133:319-330
46. Terol J, Perez-Alonso M, Frutos Rd (1995) Molecular characterization of the *zerknüllt* region of the Antennapedia complex of *D. subobscura*. Chromosoma 103:613-624
47. Frönhofer HG, Nüsslein-Volhard C (1986) Organization of anterior pattern in the *Drosophila* embryo by the maternal gene *bicoid*. Nature 324:120-125
48. Driever W, Nüsslein-Volhard C (1988) The Bicoid protein determines position in the *Drosophila* embryo in a concentration-dependent manner. Cell 54:95-104
49. Schröder R, Sander K (1993) A comparison of transplantable *bicoid* activity and partial *bicoid* homeobox sequences in several *Drosophila* and blowfly species (Calliphoridae). Roux's Arch Dev Biol 203:34-43
50. Mann RS (1995) The functional specificity of homeotic gene function. BioEssays 17:855-863
51. Gehring WJ (1994) A history of the homeobox, in Guidebook to the Homeobox Genes (Duboule D, ed.) pp 1-10. Oxford University Press, Oxford
52. Treisman J, Harris E, Wilson D, Desplan C (1992) The homeodomain: a new face for the helix-turn-helix? BioEssays 14:145-150
53. Dawes R, Dawson I, Falciani F, Tear G, Akam M (1994) *Dax*, a locust Hox gene related to *fushi-tarazu* but showing no pair-rule expression. Development 120:1561-1572
54. Brown SJ, Hilgenfeld RB, Denell RE (1994) The beetle *Tribolium castaneum* has a *fushi tarazu* homolog expressed in stripes during segmentation. Proc Natl Acad Sci USA 91:12922-12926
55. Wirz J, Fessler L, Gehring WJ (1986) Localization of the *Antennapedia* protein in the *Drosophila* embryo and imaginal discs. EMBO J 5:3327-3334
56. Pultz MA, Diederich RJ, Cribbs DL, Kaufman TC (1988) The *proboscipedia* locus of the Antennapedia complex: a molecular and genetic analysis. Genes Dev 2:901-920
57. Diederich RJ, Merrill VKL, Pultz MA, Kaufman TC (1989) Isolation, structure and expression of *labial*, a homeotic gene of the Antennapedia complex involved in *Drosophila* head development. Genes Dev 3:399-414
58. Gindhart JG, King AN, Kaufman TC (1995) Characterization of the *cis*-regulatory region of the *Drosophila* homeotic gene *Sex combs reduced*. Genetics 139:781-795
59. Hoppler S, Bienz M (1994) Specification of a single cell type by a *Drosophila* homeotic gene. Cell 76:689-702
60. Martin CH, Mayeda CA, Davis CA, Ericsson CL, Knafels JD, Mathog DR, Celniker SE, Lewis EB, Palazzolo MJ (1995) Complete sequence of the bithorax complex of *Drosophila*. Proc Natl Acad Sci USA 92:8398-8402
61. Duncan I (1987) The bithorax complex. Annu Rev Genet 21:285-319
62. Hiromi Y, Kuroiwa A, Gehring WJ (1985) Control elements of the *Drosophila* segmentation gene *fushi tarazu*. Cell 43:603-613
63. Izpisúa-Belmonte J-C, Falkenstein H, Dollé P, Renucci A, Duboule D (1991) Murine genes related to the *Drosophila AbdB* homeotic gene are sequentially expressed during development of the posterior part of the body. EMBO J 10:2279-2289
64. Buttgereit D, Renkawitzpohl R (1993) Expression of beta-1 tubulin (beta-tub56D) in *Drosophila* testis stem-cells is regulated by a short upstream sequence while intron elements guide expression in somatic-cells. Molec Gen Genet 241:263-270

65. Edgar BA, Lehman DA, O'Farrell PH (1994) Transcriptional regulation of string (cdc25): a link between developmental programming and the cell cycle. Development 120:3131-3143
66. Hanks M, Wurst W, Anson-Cartwright L, Auerbach AA, Joyner AL (1995) Rescue of the *En-1* mutant phenotype by replacement of *En-1* with *En-2*. Science 269:679-682
67. Gao D, Kimble J (1995) APX-1 can substitute for its homolog LAG-2 to direct cell interactions throughout *Caenorhabditis elegans* development. Proc Natl Acad Sci USA 92:9839-9842
68. Li X, Noll M (1994) Evolution of distinct developmental functions of three *Drosophila* genes by acquisition of different *cis*-regulatory elements. Nature 367:83-87
69. Rothe M, Pehl M, Taubert H, Jäckle H (1992) Loss of gene function through rapid mitotic cycles in the *Drosophila* embryo. Nature 359:156-159
70. Mahaffey JW, Diederich RJ, Kaufman TC (1989) Novel patterns of homeotic protein accumulation in the head of the *Drosophila* embryo. Development 105:167-174
71. Beeman R, Stuart JJ, Haas MS, Denell RE (1989) Genetic analysis of the homeotic gene complex (HOM-C) in the beetle *Tribolium castaneum*. Dev Biol 133:196-209
72. Beeman RW, Stuart JJ, Brown SJ, Denell RE (1993) Structure and function of the homeotic gene complex (HOM-C) in the beetle, *Tribolium castaneum*. BioEssays 15:439-444
73. Fleig R, Walldorf U, Gehring WJ, Sander K (1988) In situ localization of the transcripts of a homeobox gene in the honeybee *Apis mellifera* L. (Hymenoptera). Roux's Arch Dev Biol 197:269-274
74. Fleig R, Walldorf U, Gehring WJ, Sander K (1992) Development of the *Deformed* protein pattern in the embryo of the honeybee *Apis mellifera* L. (Hymenoptera). Roux's Arch Dev Biol 201:235-242
75. Tear G (1990) A molecular analysis of embryonic development in the short germ band insect *Schistocerca gregaria*. PhD thesis. Univ. Cambridge
76. Hayward DC, Patel NH, Rehm EJ, Goodman CS, Ball EE (1995) Sequence and expression of a grasshopper *Antennapedia*: comparison to *Drosophila*. Dev Biol 172:452-465
77. White RAH, Wilcox M (1985) Distribution of Ultrabithorax proteins in *Drosophila*. EMBO J 4:2035-2043
78. Ueno K, Hui C-C, Fukuta M, Suzuki Y (1992) Molecular analysis of the deletion mutants in the E homeotic complex of the silkworm *Bombyx mori*. Development 114:555-563
79. Kelsh R, Weinzierl ROJ, White RAH, Akam M (1994) Homeotic gene expression in the locust Schistocerca: an antibody that detects conserved epitopes in *Ultrabithorax* and *Abdominal-A* proteins. Dev Genet 15:19-31
80. Nagy LM, Booker R, Riddiford LM (1991) Isolation and embryonic expression of an *abdominal-A*-like gene from the lepidopteran, *Manduca sexta*. Development 112:119-129
81. Stuart JJ, Brown SJ, Beeman RW, Denell RE (1993) The *Tribolium* homeotic gene *Abdominal* is homologous to *abdominal-A* of the *Drosophila* bithorax complex. Development 117:233-243
82. Tear G, Akam M, Martinez-Arias A (1990) Isolation of an *abdominal-A* gene from the locust *Schistocerca gregaria* and its expression during early embryogenesis. Development 110:915-925
83. Zhao JJ, Lazzarini RA, Pick L (1993) The mouse *Hox-1.3* gene is functionally equivalent to the *Drosophila Sex combs reduced* gene. Genes Dev 7:343-354
84. Sham MH, Hunt P, Nonchev S, Papalopulu N, Graham A, Boncinelli E, Krumlauf R (1992) Analysis of the murine *Hox-2.7* gene: conserved alternative transcripts with differential distributions in the nervous system and the potential for shared regulatory elements. EMBO J 11:1825-1836
85. Condie BG, Capecci MR (1993) Mice homozygous for a targeted disruption of *Hoxd-3* (*Hox-4.1*) exhibit anterior transformations of the first and second cervical vertebrae, the atlas and the axis. Development 119:579-595
86. Driever W, Nüsslein-Volhard C (1988) A gradient of *bicoid* protein in *Drosophila* embryos. Cell 54:83-93
87. Seeger MA, Kaufman TC (1990) Molecular analysis of the *bicoid* gene from *Drosophila pseudoobscura*: identification of conserved domains within coding and non-coding regions of the *bicoid* mRNA. EMBO J 9:2977-2987
88. Sommer R, Tautz D (1991) Segmentation gene expression in the housefly *Musca domestica*. Development 113:419-430
89. Carroll SB, Scott MP (1985) Localization of the *fushi tarazu* protein during *Drosophila* embryogenesis. Cell 43:47-57
90. Jost W, Yu Y, Pick L, Preiss A, Maier D (1995) Structure and regulation of the *fushi tarazu* gene from *Drosophila hydei*. Roux's Arch Dev Biol 205:160-170
91. Macdonald PM, Ingham P, Struhl G (1986) Isolation, structure and expression of *even-skipped*: a second pair-rule gene of *Drosophila* containing a homeobox. Cell 47:721-734
92. Frasch M, Hoey T, Rushlow C, Doyle HJ, Levine M (1987) Characterization and localization of the *even-skipped* protein of *Drosophila*. EMBO J 6:749-759
93. Nüsslein-Volhard C, Kluding H, Jürgens G (1985) Genes affecting the segmental subdivision of the *Drosophila* embryo. Cold Spring Harb Symp Quant Biol 50:145-154
94. Doe CQ, Smouse D, and Goodman CS (1988) Control of neuronal fate by the Drosophila segmentation gene *even-skipped*. Nature 333:376-378
95. Patel NH, Condron BG, Zinn K (1994) Pair-rule expression patterns of *even-skipped* are found in both short- and long-germ beetles. Nature 367:429-434
96. Patel NH, Ball EE, Goodman CS (1992) Changing role of *even-skipped* during the evolution of insect pattern formation. Nature 357:339-342
97. Spyropoulos DD, Capecchi MR (1994) Targeted disruption of the *even-skipped* gene, *evx1*, causes early postimplantation lethality of the mouse conceptus. Genes Dev 8:1949-1961
98. Ahringer J (1996) Posterior patterning by the *C. elegans even-skipped* homolog *vab-7*. Genes Dev, in press
99. Mlodzik M, Fjose A, Gehring WJ (1985) Isolation of *caudal*, a *Drosophila* homeobox-containing gene with maternal expression whose transcripts form a concentration gradient at the pre-blastoderm stage. EMBO J 4:2961-2969
100. Macdonald PM, Struhl G (1986) A molecular gradient in early *Drosophila* embryos and its role in specifying the body pattern. Nature 324:537-545
101. Xu X, Xu PX, Suzuki Y (1994) A maternal homeobox gene, *Bombyx caudal*, forms both mRNA and protein concentration gradients spanning anteroposterior axis gastrulation. Development 120:277-285
102. Meyer BI, Gruss P (1993) Mouse *Cdx-1* expression during gastrulation. Development 117:191-203
103. Subramanian V, Meyer BI, Gruss P (1995) Disruption of the murine homeobox gene *Cdx-1* affects axial skeletal identities by altering the mesodermal expression domains of Hox genes. Cell 83:641-653
104. Gamer LW, Wright CVE (1993) Murine *Cdx-4* bears striking similarities to the *Drosophila Caudal* gene in its homeodomain sequence and early expression pattern. Mech Dev 43:71-81
105. Frumkin A, Pillemer G, Haffner R, Tarcic N, Gruenbaum Y, Fainsod A (1994) A role for *CdxA* in gut closure and intestinal epithelia differentiation. Development 120:253-263
106. Waring DA, Kenyon C (1990) Selective silencing of cell communication influences anteroposterior pattern formation in *C. elegans*. Cell 60:123-131
107. Dalton D, Chadwick R. McGinnis W (1989) Expression and embryonic function of *empty spiracles*: a *Drosophila* homeobox gene with two patterning functions on the anterior-posterior axis of the embryo. Genes Dev 3:1940-1956

108. Cohen SM, Jürgens G (1990) Mediation of *Drosophila* head development by gap-like segmentation genes. Nature 346:482-485
109. Simeone A, Acampora D, Gulisano M, Stornaiuolo A, Boncinelli E (1992) Nested expression domains of four homeobox genes in developing rostral brain. Nature 358:687-690
110. Kastury K, Druck T, Huebner K, Barletta C, Acampora D, Simeone A, Faiella A, Boncinelli E (1994) Chromosome locations of human *EMX* and *OTX* genes. Genomics 22:41-45
111. Finkelstein R, Smouse D, Capaci TM, Spradling AC, Perrimon N (1990) The *orthodenticle* gene encodes a novel homeodomain protein involved in the development of the *Drosophila* nervous system and ocellar visual structures. Genes Dev 4:1516-1527
112. Matsuo I, Kuratani S, Kimura C, Takeda N, Aizawa S (1995) Mouse *Otx2* functions in the formation and patterning of rostral head. Genes Dev 9:2646-2658
113. Chouinard S, Kaufman TC (1991) Control of expression of the homeotic *labial* (*lab*) locus of *Drosophila melanogaster*: evidence for both positive and negative autogenous regulation. Development 113:1267-1280
114. Kapoun AM, Kaufman TC (1995) A functional analysis of 5', intronic and promoter regions of the homeotic gene *proboscipedia* in *Drosophila melanogaster*. Development 121:2127-2141
115. Bergson C, McGinnis W (1990) An autoregulatory enhancer element of the *Drosophila* homeotic gene *Deformed*. EMBO J 9:4287-4297
116. Pick L, Schier A, Affolter M, Schmidt-Glenewinkel T, Gehring WJ (1990) Analysis of the *ftz* upstream element: germ layer-specific enhancers are independently autoregulated. Genes Dev 4:1224-1239
117. Boulet AM, Scott MP (1988) Control elements of the P2 promoter of the *Antennapedia* gene. Genes Dev 2:1600-1614
118. Zink B, Engstrom Y, Gehring WJ, Paro R (1991) Direct interaction of the *Polycomb* protein with *Antennapedia* regulatory sequences in polytene chromosomes of *Drosophila melanogaster*. EMBO J 10:153-162
119. Kornfeld K, Saint RB, Beachy PA, Harte PJ, Peattie DA, Hogness DS (1989) Structure and expression of a family of *Ultrabithorax* mRNAs generated by alternative splicing and polyadenylation in *Drosophila*. Genes Dev 3:243-258
120. Qian S, Capovilla M, Pirrotta V (1991) The *bx* region enhancer, a distant *cis*-control element of the *Drosophila Ubx* gene and its regulation by *hunchback* and other segmentation genes. EMBO J 10:1415-1425
121. Müller J, Bienz M (1991) Long range repression conferring boundaries of *Ultrabithorax* expression in the *Drosophila* embryo. EMBO J 10:3147-3155
122. Chan C-S, Rastelli L, Pirrotta V (1994) A *Plycomb* response element in the *Ubx* gene that determines an epigenetically inherited state of repression. EMBO J 13:2553-2564
123. Christen B, Bienz M (1992) A *cis*-element mediating *Ultrabithorax* autoregulation in the central nervous system. Mech Dev 39:73-80
124. Thuringer F, Cohen SM, Bienz M (1993) Dissection of an indirect autoregulatory response of a homeotic *Drosophila* gene. EMBO J 12:2419-2430
125. Simon J, Chiang A, Bender W, Shimell MJ, O'Connor M (1993) Elements of the *Drosophila* Bithorax Complex that mediate repression by *Polycomb* group products. Dev Biol 158:131-144
126. Celniker SE, Keelan DJ, Lewis EB (1989) The molecular genetics of the bithorax complex of *Drosophila*: characterization of the products of the *Abdominal-B* domain. Genes Dev 3:1424-1436
127. Busturia A, Bienz M (1993) Silencers in *Abdominal-B*, a homeotic *Drosophila* gene. EMBO J 12:1415-1425