

Functional evolution of Hox proteins in arthropods

Michel Vervoort

Summary

It is presumed that the evolution of morphological diversity in animals and plants is driven by changes in the developmental processes that govern morphology, hence basically by changes in the function and/or expression of a defined set of genes that control these processes. A large body of evidence has suggested that changes in developmental gene regulation are the predominant mechanisms that sustain morphological evolution, being much more important than the evolution of the primary sequences and functions of proteins. Recent reports^(1,2) challenge this idea by highlighting functional evolution of Hox proteins during the evolutionary history of arthropods. *BioEssays* 24:775–779, 2002.

© 2002 Wiley Periodicals, Inc.

Introduction

Hoxb6 (formerly *Hox-2.2*) and *Hoxd4* (formerly *Hox-4.2*) are mouse genes that are closely related to the *Drosophila* genes, *Antennapedia* (*Antp*) and *Deformed* (*Dfd*), respectively. In 1990, McGinnis and co-workers showed that the ubiquitous expression of these genes in *Drosophila* embryos mimics the effects of the ubiquitous expression of their *Drosophila* counterparts.^(3,4) Since this pioneering work, a large number of proteins from a large variety of species have been expressed in *Drosophila* and, in many cases, the expression of these proteins gives essentially identical phenotypical effects to those induced by similar expression of their *Drosophila* homologues. The induction of ectopic eyes in *Drosophila*, following ectopic expression of *Pax6* genes from various species, is a striking illustration of the outcome of this type of experiment.⁽⁵⁾ In many cases, the non-*Drosophila* protein, if appropriately expressed, can even compensate for a loss-of-function mutation in its *Drosophila* homologue, reverting the phenotype due to the mutation to the wild-type phenotype. This often occurs even when the homologues obviously perform different developmental functions in their respective species as in the following

example. The genes belonging to the *Achaete–Scute* (*AS*) complex encode basic helix–loop–helix proteins that control early events of neurogenesis in *Drosophila*.⁽⁶⁾ Homologues of these genes are known in a wide variety of species, including hydra. The single hydra *AS* gene (known as *CNASH*) is involved in the differentiation of cnidocytes but not in the formation of neural cells.⁽⁷⁾ Yet, the forced expression of *CNASH* in *Drosophila* rescues the neural phenotype due to the loss-of-function of *AS* genes.⁽⁷⁾

These and other similar results led to the idea that the “molecular function” of the proteins (i.e. their biochemical properties, as for example DNA-binding specificity or the ability to interact with other proteins) does not vary much. The main driving force of evolutionary changes would then lie in changes in gene regulation (“regulatory evolution”), which would result in well-conserved proteins functioning in different cellular contexts, due to changes in the regulation of their own expression and/or in that of the genes or proteins that they interact with.⁽⁸⁾ Two recent articles^(1,2) show, on the contrary, that evolutionary changes can also be induced by the functional evolution of the proteins themselves. The authors develop their analysis, almost ironically, on the prototype of evolutionary conserved proteins, the Hox family.

The *Ultrabithorax* (*Ubx*) gene and the formation of appendages in arthropods

The homeobox-containing Hox genes control patterning along the anteroposterior axis in a wide variety of organisms throughout the animal kingdom. These genes are clustered in complexes, highly conserved through evolution. Structurally related Hox genes are located in equivalent positions within Hox complexes of *Drosophila* and mouse, have equivalent functional properties, and are expressed in equivalent positions along the anteroposterior axis of these animals.⁽⁸⁾ This suggested that the present-day Hox complexes arose from a complex already present in the last common ancestor of both *Drosophila* and mouse, i.e. given the widely accepted animal phylogeny,⁽⁹⁾ that of all animals with a bilateral symmetry. This ancestral complex would have owned at least seven different Hox genes,⁽¹⁰⁾ already involved in some sort of patterning along the anteroposterior axis.⁽⁸⁾ Given the dramatic phenotypic consequences of loss- or gain-of-function mutations of Hox genes (homeotic transformations), they are ideal candidates for genes whose naturally occurring modifications may drive the diversification of body plans.^(8,11) This can be well exemplified

Funding agency: The CNRS, the Institut Francais de la Biodiversite, and the Universite Paris-Sud.

Correspondence to: Michel Vervoort, Evolution et Developpement des protostomiens, Centre de Genetique Moleculaire, UPR 2167 CNRS, 1, av. de la terrasse, 91198 Gif-sur-Yvette Cedex, France.

E-mail: vervoort@cgm.cnrs-gif.fr

DOI 10.1002/bies.10146

Published online in Wiley InterScience (www.interscience.wiley.com).

by the *Drosophila Hox* gene *Ultrabithorax (Ubx)* and its homologues in other arthropods.⁽⁸⁾

Among numerous functions, *Ubx* is responsible in *Drosophila* for the fact that the third thoracic segment (T3), in which *Ubx* is expressed, develops a highly modified appendage (the haltere) instead of the wing produced by the second thoracic segment (T2), which lacks *Ubx* expression. In addition, in conjunction with another Hox gene, *Abdominal-A (Abd-A)*, *Ubx* results in the abdomen of the fly lacking limbs. These are interesting functions in an evolutive perspective, as most insects, unlike *Drosophila* (and related flies, the dipterans), have wings on both T2 and T3, and most non-insect arthropods have appendages on most or even all segments (crustaceans and myriapods are good examples of this). It is tempting to think that the transition from four-winged to two-winged insects and the restriction of appendage formation to head and thoracic segments that has occurred at some point during arthropod evolution may have involved some changes in *Ubx* function. This appears not only to be tempting but also to be true.

Let's first summarize the action of *Ubx* in *Drosophila*. *Ubx* encodes a transcription factor whose function is basically to regulate gene expression. In order to understand how *Ubx* functions, statements such as *Ubx* promotes haltere formation or represses limb development should hence be translated into statements such as *Ubx* when present in particular cells activates this set of target genes and represses that other set of genes. Fortunately, this translation has been partially made and the effects of *Ubx* appear to be mainly negative in the concerned processes, i.e. through the repression of the expression of several genes.^(12,13) In particular, the effect on abdominal limb development seems to be mainly mediated through the repression of the expression of the *Distal-less (Dll)* homeobox gene. This repression is induced through binding of *Ubx* and *Abd-A* on the enhancer of *Dll* that drives its expression in the appendages primordia of early embryos.⁽¹²⁾ In the absence of *Ubx* and *Abd-A*, such as in the thoracic segments of wild-type flies or in the abdomen of *Ubx* and *Abd-A* mutants, *Dll* is expressed, allowing the formation of primordia and eventually the development of limbs. In the presence of *Ubx* and/or *Abd-A*, such as in the abdomen of wild-type flies or in the thorax of flies where *Ubx* or *Abd-A* are ubiquitous expressed, *Dll* is not expressed, the primordia do not form and the limbs are lacking. Similarly but in a more-complex manner, the presence of *Ubx* in the developing primordia of the T3 segment of wild-type flies, or in T2 segments following forced expression of *Ubx*, prevents or modifies the expression of several patterning genes required for proper wing development (such as *wingless*, *D-SRF* and *achaete*), therefore giving rise to a highly reduced appendage, the haltere.⁽¹³⁾ In the wild-type T2 segment (or in a T3 segment from an *Ubx* mutant fly), these genes are not repressed and a true wing forms.

What is the situation in four-winged insects and arthropods with abdominal legs? An *Ubx* gene has been found in all tested

arthropods to date and appears to be expressed in a very similar pattern to their *Drosophila* counterpart.^(14–19) However, *Ubx* is co-expressed with *Dll* in embryos of several arthropod species, such as crustaceans,⁽¹⁵⁾ myriapods,⁽¹⁶⁾ “lower” insects (collembolans),⁽¹⁷⁾ and even in “higher” insects such as the coleopteran *Tribolium castaneum*.^(17,18) As a consequence, in these different species, some appendages form from the *Ubx*-expressing part of the body. In butterflies, which are quite closely related to dipterans, yet are four-winged insects, a true hindwing develops from a primordium that continuously expresses *Ubx*.⁽¹⁴⁾ In the butterfly hindwing primordium, *Ubx* appears to be unable to repress the expression of genes such as *wingless*, *D-SRF*, and *achaete*, unlike it does in *Drosophila*.⁽²⁰⁾ As the evolution transitions are likely to be from four- to two-winged insects and from limbs on every trunk segments to limbless abdomen, it appears that, during the evolution of arthropods, *Ubx* has gained the ability to repress the aforesaid target genes.

Functional evolution of *Ubx* in arthropods

How did this acquisition occur? In theory, this can be achieved by changes in the *cis*-regulatory elements of the target genes and/or in the functional properties of the *trans*-acting regulatory protein (in this case, *Ubx* itself). The second possibility can be quite easily tested by expressing *Ubx* from various species in *Drosophila* and monitoring whether it induces the same effects as *Drosophila Ubx*. Three different *Ubx* proteins have been used in such tests, from the onychophoran species, *Akanthokara kaputensis* (the onychophora is a sister phylum of arthropods with a simple body plan, in particular with simple unjointed limbs on every segments),^(1,21) from the crustacean *Artemia franciscana*,⁽²⁾ and from the coleopteran *Tribolium castaneum*.⁽¹⁾ The latter mimics all the effects of *Drosophila Ubx* during embryogenesis, in particular the ability to repress *Dll* expression and therefore limb formation (Fig. 1A).⁽¹⁾ On the contrary, the onychophoran and crustacean *Ubx* are unable to repress *Dll* expression and to prevent limb formation, although they differ in their ability to promote the transformation of thoracic segments to an abdominal fate^(1,2,21) (Fig. 1G,H). From this set of experiments, one can conclude that the *Ubx* protein has gained, both in crustaceans and insects but not onychophorans, molecular functions that allow it to promote abdominal fate, and that *Ubx* protein has gained molecular functions to repress limb formation in insects only. In contrast, the ability of *Drosophila Ubx* to repress some wing-patterning genes, hence leading to haltere formation, appears not to be related to modifications in the function of the *Ubx* protein; indeed, the onychophoran *Ubx* protein, when expressed in the developing wing primordium, is able to repress *D-SRF* and to promote a wing-to-haltere transformation.⁽²¹⁾ Therefore, the ability to repress wing development is not related to modifications of *Ubx* but most probably to changes in the regulatory

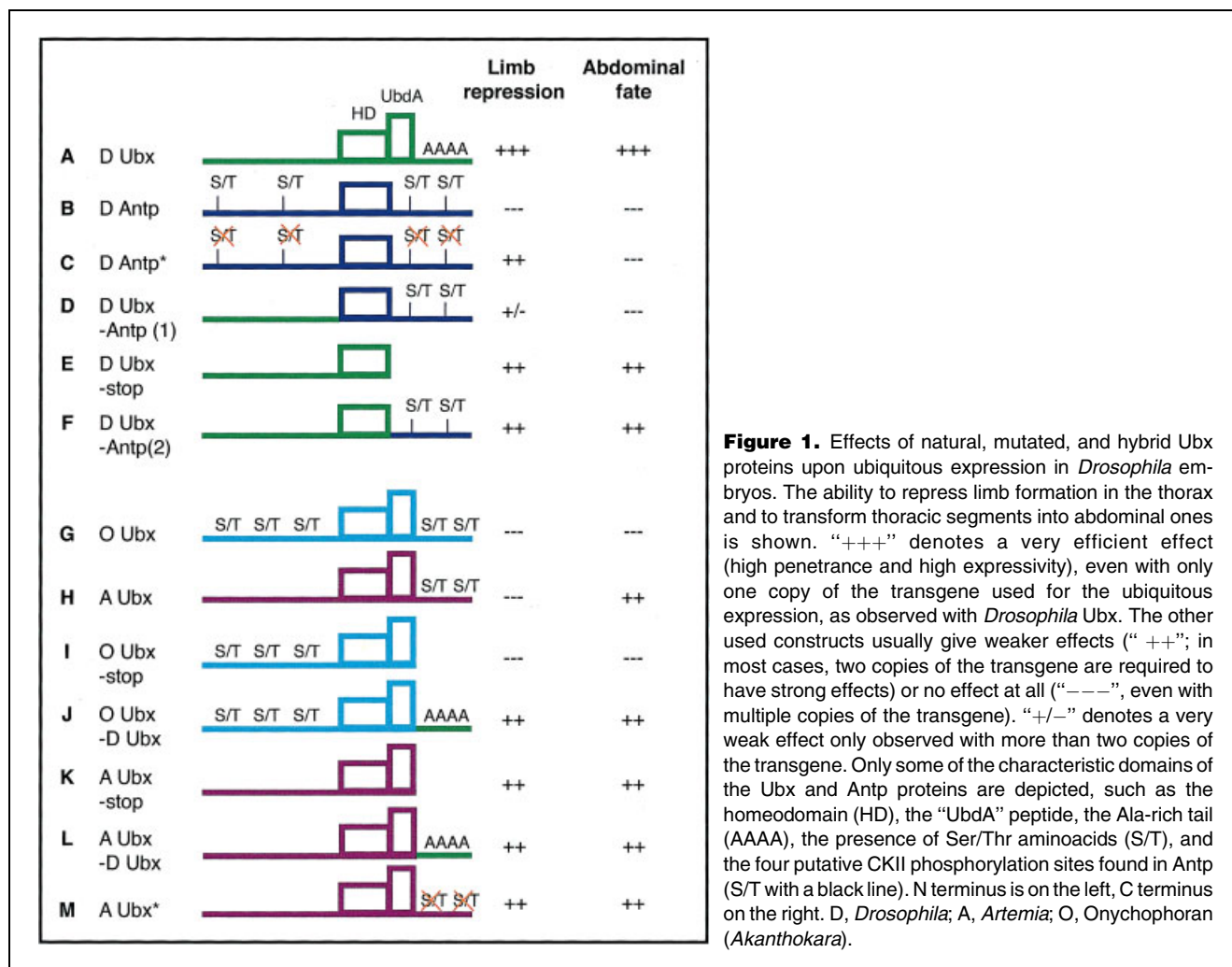


Figure 1. Effects of natural, mutated, and hybrid Ubx proteins upon ubiquitous expression in *Drosophila* embryos. The ability to repress limb formation in the thorax and to transform thoracic segments into abdominal ones is shown. “+++” denotes a very efficient effect (high penetrance and high expressivity), even with only one copy of the transgene used for the ubiquitous expression, as observed with *Drosophila* Ubx. The other used constructs usually give weaker effects (“++”; in most cases, two copies of the transgene are required to have strong effects) or no effect at all (“---”, even with multiple copies of the transgene). “+/-” denotes a very weak effect only observed with more than two copies of the transgene. Only some of the characteristic domains of the Ubx and Antp proteins are depicted, such as the homeodomain (HD), the “UbdA” peptide, the Ala-rich tail (AAAA), the presence of Ser/Thr aminoacids (S/T), and the four putative CKII phosphorylation sites found in Antp (S/T with a black line). N terminus is on the left, C terminus on the right. D, *Drosophila*; A, *Artemia*; O, Onychophoran (*Akanthokara*).

region of target genes such as *D-SRF*. This hypothesis has, however, still to be experimentally tested.

Let's now focus on the evolution of the Ubx protein. Truncated, mutant and hybrid proteins have been used to investigate the differences between insect, crustacean, and onychophoran Ubx proteins.^(1,2,21) Similar analyses were conducted in the early 1990s into the functional differences between two related Hox proteins from *Drosophila*, Ubx and Antp (unlike Ubx, Antp is unable to repress limb formation and to promote thoracic-to-abdominal fate transformations; Fig. 1A,B).^(22–24) The more relevant data from these experiments are depicted in Fig. 1 and lead to the following conclusions.

(1) The ability of *Drosophila* Ubx to promote abdominal fate is a specific function of Ubx (and Abd-A) proteins. The specificity lies in its homeodomain and a few aminoacids directly C-terminal to the homeodomain, the so-called UbdA peptide

(present in Ubx, Abd-A and their homologues in various protostome species but not in other Hox proteins).⁽¹⁰⁾

(2) In contrast, the ability to repress *Dll* expression and limb formation appears to be related to more general properties of Hox proteins such as Ubx and Antp. Antp is unable to repress limb formation but this is due to its phosphorylation, in particular by Casein Kinase II (CKII). Mutations in the prospective CKII phosphorylation sites of Antp is sufficient to transform this protein into a limb-repressive proteins; however, it is less efficient than Ubx (Fig. 1C). These sites are absent in *Drosophila* Ubx but several serine (Ser) and threonine (Thr) aminoacids as well as two putative CKII consensus sites are present in the C-terminal part of *Artemia* Ubx. Suppressing this C-terminal part (Fig. 1K), replacing it with the corresponding domain from *Drosophila* (Fig. 1L), or even simply converting the Ser and Thr aminoacids into alanine (Ala) (Fig. 1M) is sufficient to confer a limb-repressive activity to *Artemia* Ubx.

- (3) The C-terminal part of *Drosophila* Ubx, which contains many Ala residues, helps in the repression of limb development. Indeed, the replacement of the C-terminal part of *Akanthokara* Ubx by that of *Drosophila* is sufficient to produce a potent limb-repressive protein (Fig. 1J). Cell culture analysis indicates that this Ala-rich C-terminal domain can act as a repression domain.⁽¹⁾ It is, however, not strictly required, for *Dll* repression, as a *Drosophila* Ubx protein lacking this domain still has a limb-repressive activity, albeit reduced in efficiency (Fig. 1E).
- (4) In addition to the presence of the Ala-rich C-terminal tail and the absence of phosphorylation sites, specificities in the homeodomain also has some importance in the limb-repressing activity of *Drosophila* Ubx. Indeed, a hybrid protein made of the N-terminal part of Ubx followed by the homeodomain and the C-terminal domain of Antp is a very poor repressor of limb development (Fig. 1D). This can be best explained by the phosphorylation of the proteins due to the presence of CKII consensus sites in the C-terminal part of Antp. However, a hybrid protein made of the N-terminal part and homeodomain of Ubx followed by the C-terminal domain of Antp represses limb formation much more efficiently; in fact, the repression is comparable to a truncated Ubx that lacks any C-terminal domain (Fig. 1E,F).

These results suggest that the ability of *Drosophila* Ubx to repress *Dll* expression and limb formation is not directly related to one particular domain of the protein but rather involves several properties of the proteins, such as the absence of some phosphorylation sites, the presence of an Ala-rich repression domain and some specificities in the homeodomain.

Conclusions

How was the ability to repress limb development gained during arthropod evolution? My belief is that there was not really an “invention” of a limb-repressive domain that would have been first conditional (in non-insect arthropods) and then become

constitutive (in insects), as it has been proposed.^(1,2,25,26) I would rather suggest that the ability of Ubx to repress limb development in insect is related to a property of Ubx proteins that was originally devoted to functions unrelated to limb development and subject to negative regulation through phosphorylation (Fig. 2). This regulation may have been involved in modulating Ubx functions in relation to levels of particular kinases such as CKII and even in inhibiting Ubx activity in cellular contexts characterized by high levels of those kinases. Mutations that would have reduced this regulation may have been selected as giving rise to a more efficient Ubx in those original functions or an Ubx with new abilities (such as acting in previously “forbidden” cellular contexts). I would suggest that the acquisition of the ability to promote an abdominal fate has been acquired by such mutations that partially free the Ubx protein from negative regulation. This may have occurred after the divergence between onychophorans and arthropods. Indeed, the *Akanthokara* Ubx when expressed in *Drosophila* is not able to promote abdominal fate, in contrast to *Drosophila* and *Artemia* Ubx.^(1,2) *Akanthokara* Ubx has many Ser and Thr aminoacids not only in its C-terminal domain (as other non-insect Ubx) but also throughout its sequence (unlike arthropod Ubx). These sites may prevent *Akanthokara* Ubx functioning as its arthropod counterparts.

Later in the evolution of arthropods (Fig. 2), after the divergence between insects and other arthropods, further mutations of the phosphorylation sites may have completely abolished the negative regulation, giving rise to an Ubx protein now able to repress *Dll* expression and limb formation. In my view, this means that Hox response elements were present in the regulatory region of *Dll* and some other genes involved in limb development, well before these genes became under Ubx regulation. These elements may have been the target of other Hox proteins (such as Antp) and/or involved in regulating expression in tissues other than the limbs (such as the brain). These elements would not have mediated, however, a regulation by Ubx in the limb primordia, due to the phosphorylation of Ubx in this particular cell context. The loss of this negative

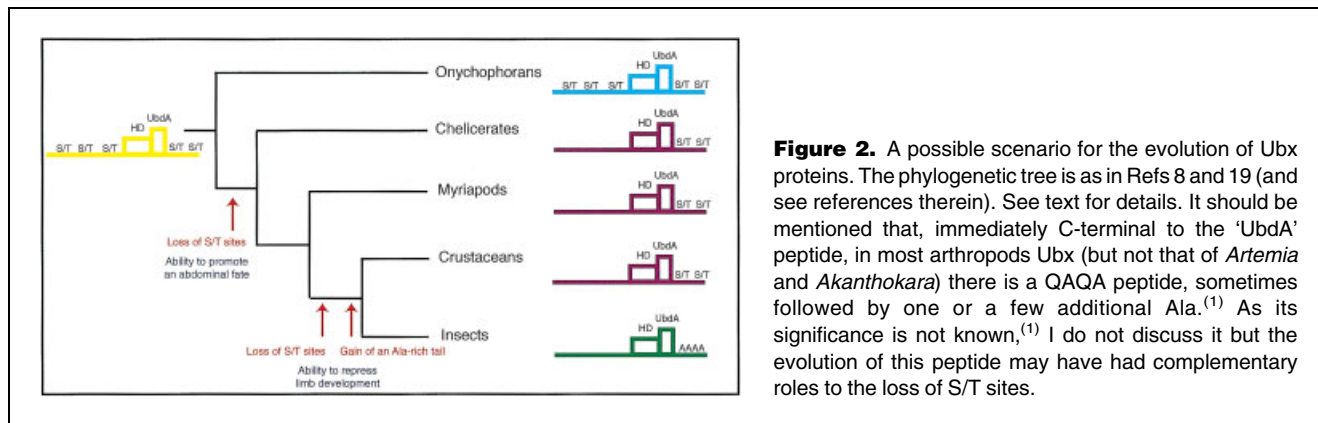


Figure 2. A possible scenario for the evolution of Ubx proteins. The phylogenetic tree is as in Refs 8 and 19 (and see references therein). See text for details. It should be mentioned that, immediately C-terminal to the ‘UbxA’ peptide, in most arthropods Ubx (but not that of *Artemia* and *Akanthokara*) there is a QAQA peptide, sometimes followed by one or a few additional Ala.⁽¹⁾ As its significance is not known,⁽¹⁾ I do not discuss it but the evolution of this peptide may have had complementary roles to the loss of S/T sites.

regulation may have therefore led to a limb-repressive Ubx. The subsequent or concomitant addition of Ala aminoacids in the C-terminal part of Ubx may then have led to the very efficient repressor of present-day insects.

Obviously, this scenario is very hypothetical and several aspects remain to be addressed in more detail. The use of Ubx from other arthropod species (myriapods for example) would be important to confirm (or refute) some of the ideas raised by the data reviewed here. There are also still some problems about Ubx function in limb development in insects. As an example, *Tribolium* Ubx when expressed in *Drosophila* mimics well the effects of its *Drosophila* homologues, including the repression of limb.⁽¹⁾ Yet, in *Tribolium*, Ubx does not repress limb formation, as cells that strongly express Ubx also express *Dll* and give rise to an appendage.⁽¹⁸⁾ Ectopic expression of Ubx proteins in *Tribolium* may help to better understand what happens in this situation. This can now be done, as ways of artificially expressing genes in arthropods other than *Drosophila* have been developed.^(27,28)

It is widely believed that changes in developmental gene regulation have played a primary role in the evolution of morphological diversity in animals.⁽⁸⁾ The two articles reviewed here give a striking demonstration that changes in protein function may also be important and may be more frequent than previously believed, although few other examples have been described.^(29–32) The two reviewed articles,^(1,2) however, not only demonstrate an evolution of the protein but also identify the possible molecular modifications that may have occurred, hence providing a possible connection between defined genetic modifications and the evolution of body plans.

Acknowledgments

I am grateful to Valerie Ledent, Benjamin Prud'homme, and Renaud de Rosa for helpful comments on the manuscript. I dedicate this review to the memory of Andre Adoutte who tragically died a few weeks ago.

References

- Galant R, Carroll SB. Evolution of a transcriptional repression domain in an insect Hox protein. *Nature* 2002;415:910–913.
- Ronshaugen M, McGinnis N, McGinnis W. Hox protein mutation and macroevolution of the insect body plan. *Nature* 2002;415:914–917.
- Malicki J, Schughart K, McGinnis W. Mouse *Hox-2.2* specifies thoracic segmental identity in *Drosophila* embryos and larvae. *Cell* 1990;63:961–967.
- McGinnis N, Kuziora MA, McGinnis W. Human *Hox-4.2* and *Drosophila deformed* encode similar regulatory specificities in *Drosophila* embryos and larvae. *Cell* 1990;63:969–976.
- Gehring WJ, Ikeo K. *Pax 6*: mastering eye morphogenesis and eye evolution. *Trends Genet* 1999;15:371–377.
- Vervoort M, Ledent V. The evolution of neural basic Helix-Loop-Helix proteins. *The ScientificWorld* 2001;1:396–426.
- Grens A, Mason E, Marsh J, Bode HR. Evolutionary conservation of a cell fate specification gene: the *Hydra achaete-scute* homolog has proneural activity in *Drosophila*. *Development* 1995;121:4027–4035.
- Carroll SB, Grenier JK, Weatherbee SD. From DNA to diversity: Molecular genetics and the evolution of animal design. London: Blackwell Science; 2001.
- Adoutte A, Balavoine G, Lartillot N, Lepinet O, Prud'homme B, de Rosa R. The new animal phylogeny: reliability and implications. *Proc Natl Acad Sci* 2000;97:4453–4456.
- de Rosa R, Grenier JK, Andreeva T, Cook CE, Adoutte A, Akam M, Carroll SB, Balavoine G. *Hox* genes in brachiopods and priapulids and protostome evolution. *Nature* 1999;399:772–776.
- Gellon G, McGinnis W. Shaping animal body plans in development and evolution by modulation of Hox expression patterns. *BioEssays* 1998;20:116–125.
- Vachon G, Cohen B, Pfeifle C, McGuffin M, Botas J, Cohen SM. Homeotic genes of the Bithorax complex repress limb development in the abdomen of the *Drosophila* embryo through the target gene *Distal-less*. *Cell* 1992;71:437–450.
- Weatherbee SD, Halder S, Kim J, Hudson Az, Carroll SB. Ultrabithorax regulates genes at several levels of the wing-patterning hierarchy to shape the development of the *Drosophila* haltere. *Genes Dev* 1998;12:1474–1482.
- Warren RW, Nagy L, Selegue J, Gates J, Carroll SB. Evolution of homeotic gene regulation and function in flies and butterflies. *Nature* 1994;372:458–461.
- Averof M, Akam M. *Hox* genes and the diversification of insect and crustacean body plans. *Nature* 1995;376:420–423.
- Grenier JK, Garber TL, Warren R, Whittington PM, Carroll SB. Evolution of the entire arthropod Hox gene set predated the origin of the onychophoran/arthropod clade. *Curr Biol* 1997;7:547–553.
- Palopoli MF, Patel NH. Evolution of the interaction between *Hox* genes and a downstream target. *Curr Biol* 1998;8:587–590.
- Lewis DL, DeCamillis M, Bennett RL. Distinct roles of the homeotic genes *Ubx* and *abdA* in beetle embryonic abdominal appendage development. *Proc Natl Acad Sci* 2000;97:4504–4509.
- Hughes CL, Kaufman TC. Exploring the myriapod body plan: expression patterns of the ten *Hox* genes in a centipede. *Development* 2002;129:1225–1238.
- Weatherbee SD, Nijhout HF, Grunert LW, Halder G, Galant R, Selegue J, Carroll SB. Ultrabithorax function in butterfly wings and the evolution of insect wing patterns. *Curr Biol* 1999;9:109–115.
- Grenier JK, Carroll SB. Functional evolution of the Ultrabithorax protein. *Proc. Natl Acad Sci USA* 2000;97:704–709.
- Mann RS, Hogness DS. Functional dissection of Ultrabithorax proteins in *D. melanogaster*. *Cell* 1990;60:597–610.
- Chan SK, Mann RS. The segment identity functions of Ultrabithorax are contained within its homeo domain and carboxy-terminal sequences. *Genes Dev* 1993;7:796–811.
- Jaffe L, Ryoo HD, Mann RS. A role for phosphorylation by casein kinase II in modulating Antennapedia activity in *Drosophila*. *Genes Dev* 1997;11:1327–1340.
- Pavlopoulos A, Averof M. Developmental evolution: Hox proteins ring the changes. *Curr Biol* 2002;12:R291–R293.
- Levine M. Hoxw insects lose their limbs. *Nature* 2002;415:848–849.
- Lewis DL, DeCamillis MA, Brunetti CR, Halder G, Kassner VA, Selegue JE, Higgs S, Carroll SB. Ectopic gene expression and homeotic transformations in arthropods using recombinant Sindbis viruses. *Curr Biol* 1999;9:1279–1287.
- Oppenheimer DI, MacNicol AM, Patel NH. Functional conservation of the *wingless-engrailed* interaction as shown by a widely applicable baculovirus misexpression system. *Curr Biol* 1999;9:1288–1296.
- Löhr U, Yussa M, Pick L. *Drosophila fushi tarazu*: a gene on the border of homeotic function. *Curr Biol* 2001;11:1403–1412.
- Alonso CR, Maxton-Kuechenmeister J, Akam M. Evolution of Ftz protein function in insects. *Curr Biol* 2001;11:1473–1478.
- Dubnau J, Struhl G. RNA recognition and translational regulation by a homeodomain protein. *Nature* 1996;379:694–699.
- Rivera-Pomar R, Niessing D, Schmidt-Ott U, Gehring WJ, Jackle H. RNA binding and translational suppression by bicoid. *Nature* 1996;379:746–749.