RESEARCH COMMENTARY

Discrimination in partnership: compartment-specific interactions of Hox proteins

K. G. GURUHARSHA[1,2] and K. VIJAYRAGHAVAN[1]

[1]*National Centre for Biological Sciences, Tata Institute of Fundamental Research, GKVK PO, Bangalore 560 065, India*
[2]*Department of Studies in Zoology, University of Mysore, Manasagangotri, Mysore 570 006, India*

A relatively small number of regulatory proteins perform multiple functions to orchestrate the development of an animal from a zygote. Many regulatory proteins are transcription factors that affect the expression of downstream genes. These factors interact with other proteins in the nucleus to regulate expression of their target genes. The main effect of loss-of-function mutations in genes encoding major regulators of development is lethality at the earliest stage where the regulatory molecule is required. Many, very important roles of regulatory factors have been inferred from these lethal phenotypes. However, the early lethality means that understanding the later roles, if any, of such genes becomes difficult. In addition, loss-of-function studies alone do not usually provide insights into the identity and role of interacting proteins critical for the final output of the regulatory signal.

Till recently, much of the focus in genomic research was on understanding gene sequences and function. These studies showed that important regulatory proteins are highly conserved in sequence and, often, in function, in organisms as diverse as mice and worms. Thus, novel genes alone do not make one organism different from another. Instead, many findings have made it clear that the difference between species very likely arises from the function of regulatory sequences that determine the time and pattern of gene expression. These patterns of gene expression define a diversity of cellular functions during development, which in turn allows, again through regulation of gene expression and by epigenetic mechanisms, further refinement of the pathways of development in each organism (see Pennisi 2004 for example).

The regulated expression of a gene is thus a consequence of a fine balance between two components: (i) *trans*-acting factors—these are transcription factors/regulatory proteins with specific affinity for each other and for target DNA sequences; and (ii) the targets of these factors or the *cis*-regulatory DNA which responds to the *trans*-acting factors by allowing or repressing gene expression. In other words, the binding of specific transcription factors to the combinatorial code of *cis*-regulatory elements brings about activation or repression of target genes (Arnone and Davidson, 1997, Ghazi and Vijay-Raghavan 2000).

Researchers have found new and efficient ways to locate regulatory regions distributed in the genome. Enhancers—stretches of DNA that harbor *cis*-regulatory elements—have been identified for some years now in model organisms such as *Drosophila* by examining how these regulatory sequences affect reporter gene expression. These studies have rapidly deciphered different mechanisms by which transcription factors and other proteins interact with regulatory DNA and with each other to guide gene activity. Reporter studies of gene enhancers with precise alterations in putative binding sites for different transcription factors in an otherwise wild-type background answer questions on gene regulation at greater resolution. These studies have revealed how regulatory pathways and their effectors ensure that a gene is turned on at the right time in the right place and, often importantly, at the right levels (Flores *et al*. 2000; Halfon *et al*. 2000). Traditionally, researchers address the functions of regulatory elements one at a time, while, more recently, a few are taking more global or bioinformatics approaches (see Halfon *et al*. 2002; Markstein *et al*. 2002, 2004; Berman *et al*. 2004 and references therein; Schroeder *et al*. 2004). Many such genomewide approaches are reviewed in a special issue of *Science* 'Genes in Action' **306** (2004). One would think that reporter-gene expression in analyses of *cis*-regulatory elements is, as a result of these studies, predictable. However, not only is this not the case, but more traditional approaches still seem to yield

E-mail: gharsha@ncbs.res.in; vijay@ncbs.res.in.

Keywords. Hox proteins; segmentation genes; *cis*-elements; binding domain; repressor; gene regulation.

surprises. One such example is reported by Gebelein *et al*. (2004) in a recent issue of *Nature*.

During *Drosophila* embryogenesis, segments, each with an anterior and posterior compartment, are generated under the control of the segmentation genes, while *Hox* genes provide each segment with a unique identity (figure 1). Classically, these two processes have been thought to occur independently. Analysis of *Distalless*—a Hox target gene—using enhancer–reporter constructs by Gebelein *et al*. (2004) reveals a previously unanticipated use of compartments for gene regulation by *Hox* proteins. Their results suggest that these segmentation and Hox gene products may collaborate to directly control gene expression of many downstream target genes.

Distalless (*Dll*) is a Hox target gene required for leg development in *Drosophila* (Cohen *et al*. 1989). It is activated by Wingless (Wg)—a Wnt family morphogen—in each thoracic hemisegment in a group of cells that straddle the anterior–posterior compartment boundary (Cohen 1990; Mann 1994). A *cis*-regulatory element derived from *Dll*, called DMX, is composed of a large activator element (DMXact) and a 57-base-pair (bp) repressor element referred to as DMX-R. DMX is extensively conserved in different *Drosophila* species, and drives accurate *Dll*-like expression in the thorax (figure 2). The abdominal *Hox* genes *Ultrabithorax* (*Ubx*) and *abdominalA* (*abdA*) cooperatively bind to DMX-R along with two homeodomain cofactors, Extradenticle (Exd) and Homothorax (Hth), and directly repress *Dll* and *DMX–lacZ* in both compartments, thereby blocking leg development in the abdomen (Gebelein *et al*. 2002, 2004).

Surprisingly, mutating a specific sequence in DMX-R results in derepression only in En (engrailed)-positive posterior compartment cells, and mutating another sequence results in derepression only in En-negative anterior compartment cells. Single mutations in the Hox1, Exd or Hth sites also result in derepression, predominantly in posterior compartment cells. In contrast, deletion of the entire DMX-R results in derepression in both compartments (figure 2). The authors find that Exd/Hth/AbdA and Exd/Hth/Ubx trimers cooperatively bind to a DMX-R DNA probe. Full-length DMX-R promotes the assembly of Hox/Exd/Hth/Hox tetramers in DNA binding experiments using DMX-R DNA probes and purified proteins. Although these binding experiments show necessary abdominal Hox input for *Dll* repression, they do not explain the compartment-specific derepression exhibited by some DMX-R mutations.

In the thorax, cells expressing *Dll* initially also synthesize either En or Sloppy paired (Slp), a Forkhead (Fkh) domain factor. In the abdomen, DMX reporter constructs lacking the DMX-R domain are coexpressed with either Slp in the anterior compartment or En in the posterior compartment. This suggested a role for these compartmentalization genes in regulating *Dll*. Mutations in the En binding domain (called X5) or a consensus binding

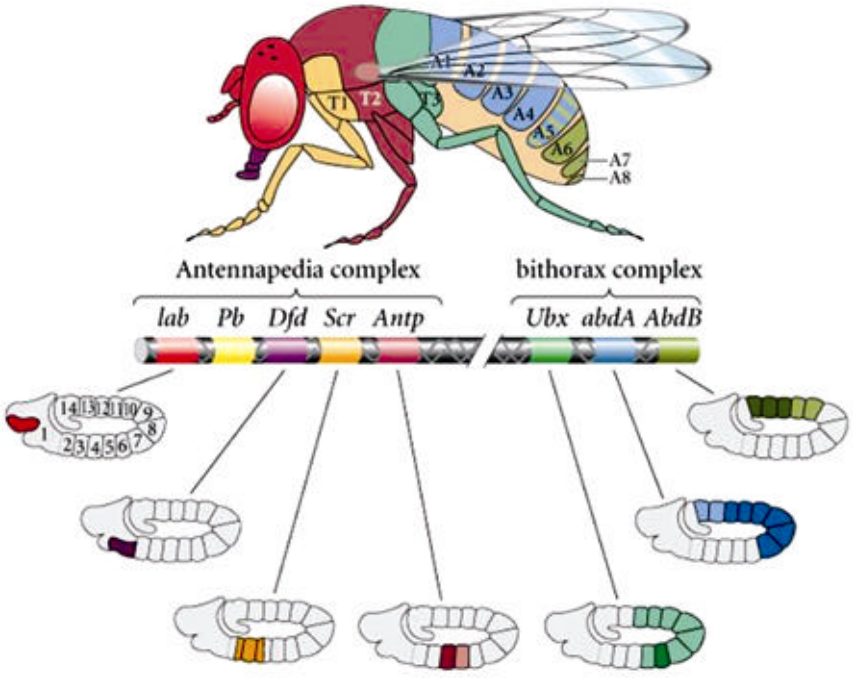

Figure 1. Schematic showing organization and expression pattern of different *Hox* genes in *Drosophila*. Taken from Gilbert's *Developmental biology* Fig. 9.27, p. 286 (after Dessain *et al*. (1992) and Kaufman *et al*. (1990)).

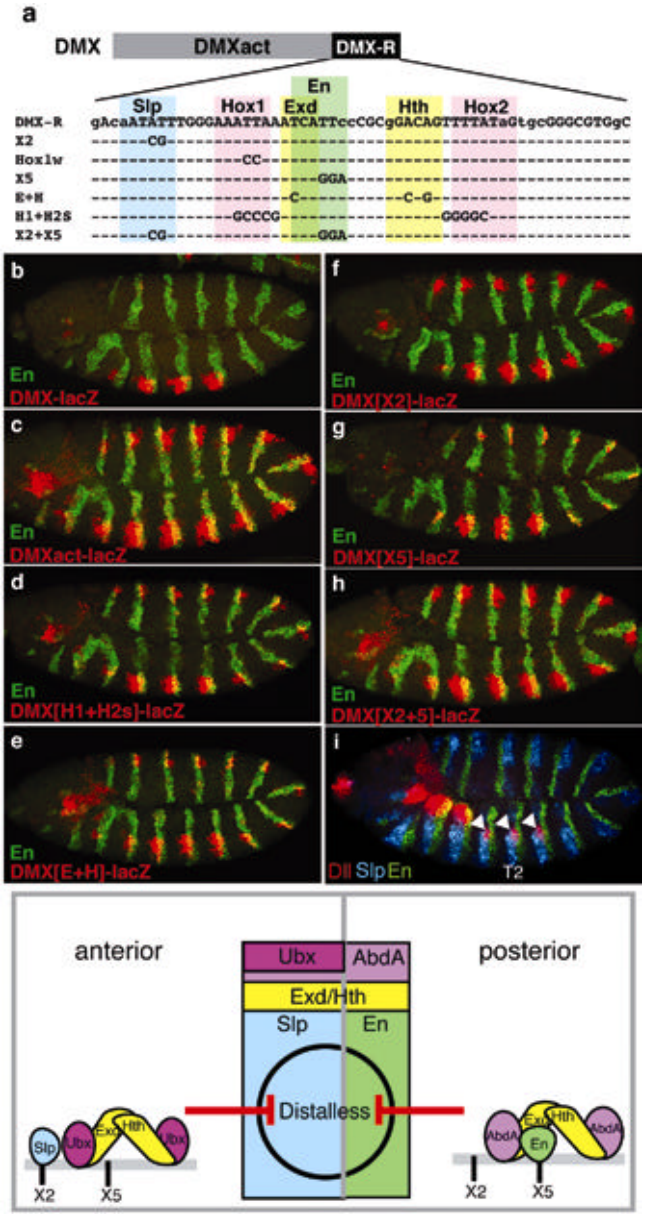

Figure 2. Analysis of *Dll cis* regulatory elements (a) Summary of mutations of DMX-Repressor sites. (b–h) Lateral views of stage-11 embryos showing lacZ production in different DMX-R animals. (i) Synthesis of Slp and En with respect to Dll. At bottom is the proposed model with a summary of the patterns of synthesis of Ubx, AbdA, Exd, Hth, Slp and En in the abdominal segments (modified from Gebelein *et al*. 2004).

site for the Fkh domain protein Slp (called X2) result in depression of reporter gene expression, respectively, in these compartments (figure 2). Binding assays reveal that En binds DMX-R along with the abdominal Hox proteins Ubx or AbdA in a highly cooperative manner, but very weakly on its own. The formation of a putative En/AbdA/Exd/Hth/AbdA complex is also predicted in these assays. Ectopic Slp represses 'wild-type' reporter-gene expression, providing strong *in vivo* support for Slp's direct role in *Dll* repression in the anterior compartments.

Based on these data, Gebelein *et al*. (2004) propose a model which suggests that in the anterior compartment Slp binds to DMX-R directly along with a Ubx/Exd/Hth/Ubx tetramer and in the posterior compartment En binds to DMX-R directly with an AbdA/Exd/Hth/AbdA tetramer (figure 2). Earlier research by other groups has shown that both Slp and En are known repressor proteins that directly bind the corepressor Groucho and, therefore, suggest a mechanism for repression. The authors also provide additional support for the model by genetic experiments using ectopic synthesis of AbdA, Ubx or synthesis of Slp with Ubx and demonstrate repression of reporter *lacZ* expression in specific compartments as predicted in the model.

The model and results presented by Gebelein *et al*. (2004) raise the question of why a compartment-specific mechanism is used by Hox factors to repress *Dll*. Alternatively, abdominal Hox proteins could have used the same set of cofactors to repress *Dll* in all abdominal cells, regardless of their compartmental origin. One possibility is that the utilization of segmentation proteins and compartment-specific mechanisms by Hox proteins may also provide additional flexibility for regulation of target genes specifically in anterior or posterior cell types. For these reasons, compartment-dependent mechanisms of gene regulation may turn out to be the general rule instead of the exception. However, it will be important to dissect other Hox-regulated elements in similar detail to assess the generality of this novel mechanism.

Acknowledgements

We wish to acknowledge Richard Mann for providing us the original figures. K.G.G. is supported by a senior research fellowship of Council of Scientific and Industrial Research, Government of India.

References

Arnone M. I. and Davidson E. H. 1997 The hardwiring of development: organization and function of genomic regulatory systems. *Development* **124**, 1851–1864.

Berman B. P., Pfeiffer B. D., Laverty T. R., Salzberg S. L., Rubin G. M., Eisen M. B. and Celniker S. E. 2004 Computational identification of developmental enhancers: conservation and function of transcription factor binding-site clusters in *Drosophila melanogaster* and *Drosophila pseudoobscura*. *Genome Biol.* **5**, R61.

Cohen S. M., Bronner G., Kuttner F., Jurgens G. and Jackle H. 1989 *Distal-less* encodes a homoeodomain protein required for limb development in *Drosophila*. *Nature* **338**, 432–434.

Cohen S. M. 1990 Specification of limb development in the *Drosophila* embryo by positional cues from segmentation genes. *Nature* **343**, 173–177.

Flores G. V., Duan H., Yan H., Nagaraj R., Fu W., Zou Y., Noll M. and Banerjee U. 2000 Combinatorial signaling in the specification of unique cell fates. *Cell* **103**, 75–85.

Gebelein B., Culi J., Ryoo H. D., Zhang W. and Mann R. S. 2002 Specificity of *Distalless* repression and limb primordia development by abdominal Hox proteins. *Dev. Cell.* **3**, 487–498.

Gebelein B., McKay D. J. and Mann R. S. 2004 Direct integration of *Hox* and segmentation gene inputs during *Drosophila* development. *Nature* **431**, 653–659.

Genes in Action. Special issue *Science* **306**. http://www.sciencemag.org/sciext/genome2004/.

Ghazi A. and VijayRaghavan K. 2000 Developmental biology. Control by combinatorial codes. *Nature* **408**, 419–420.

Halfon M. S., Carmena A., Gisselbrecht S., Sackerson C. M., Jimenez F., Baylies M. K. and Michelson A. M. 2000 Ras pathway specificity is determined by the integration of multiple signal-activated and tissue-restricted transcription factors. *Cell* **103**, 63–74.

Halfon M. S., Grad Y., Church G. M. and Michelson A. M. 2002 Computation-based discovery of related transcriptional regulatory modules and motifs using an experimentally validated combinatorial model. *Genome Res.* **12**, 1019–1028.

Mann R. 1994 engrailed-mediated repression of *Ultrabithorax* is necessary for the parasegment 6 identity in *Drosophila*. *Development* **120**, 3205–3212.

Markstein M. and Levine M. 2002 Decoding cis-regulatory DNAs in the *Drosophila* genome. *Curr. Opin. Genet. Dev.* **12**, 601–606.

Markstein M., Zinzen R., Markstein P., Yee K. P., Erives A., Stathopoulos A. and Levine M. A. 2004 Regulatory code for neurogenic gene expression in the *Drosophila* embryo. *Development* **131**, 2387–2394.

Pennisi E. 2004 Searching for the Genome's Second Code. *Science* **306**, 632–635.

Schroeder M. D., Pearce M., Fak J., Fan H., Unnerstall U., Emberly E., Rajewsky N., Siggia E. D. and Gaul U. 2004 Transcriptional control in the segmentation gene network of *Drosophila*. *PLoS Biol.* **2**, E271. 1396–1410.