

Possible functions of the homoeobox

SIR — Recent exciting findings by several laboratories have centred on the possible molecular mode of function of the *Drosophila* homoeotic gene products which appear to act as regulatory elements within the nucleus to determine the development fates of specific groups of cells¹⁻⁵. Of particular interest has been the observation of the "homoeobox", an intriguing structural element which is common not only to several genes in the *Drosophila* homoeotic gene complexes *Antennapedia* and *Bithorax*, but to other genes in *Drosophila* and in diverse other species as well³⁻¹⁰.

Although the function(s) of other genes which contain the homoeobox sequence is, at present, not known, it has been a tempting speculation that the putative protein domains encoded by the conserved homoeobox sequences may serve a common regulatory role in some aspect of pattern formation or "segmentation"; for example, Carrasco *et al.* speculate that the "homeo domain may give cells universal yet simple instructions, such as to stop dividing, to divide faster, to be considered determined, or to stop searching for new pathways of development"⁷. However, the fact that this domain has been so highly conserved in evolution suggests that it may serve a very basic conserved cellular function. We wish to point out that this function may not necessarily be directly related to the overall role of the gene product of which the homoeobox protein domain is a part.

It is a mundane but nonetheless inescapable fact of cellular life that translation occurs in the cytoplasm and that proteins with nuclear function must be transported into and/or sequestered in the nucleus. Thus, one possibility for a very basic generalized role of the homoeo domain might be in determination of the destination within the cell (the nucleus) of the gene product of which it is a part. There are, of course, other possibilities; for instance this domain might represent a generalized binding region for nucleic acid or for acidic proteins. The idea that the homoeo domain might simply represent a nuclear tag is generated by the recent work of Kalderon *et al.*¹¹, part of whose data so nicely decorated the cover of the 6-11 September issue of *Nature*. They showed that a region containing a run of five basic amino acids is solely responsible for nuclear localization of the SV40 large-T antigen; indeed, they found that replacement of the first Lys residue with a noncharged amino acid is sufficient to cause this protein to remain entirely cytoplasmic. One cannot help but compare this striking result with the fact that the key conserved feature of the 60 amino acid homoeobox protein sequences are three domains in which similar short runs of basic amino acids occur.

Consistent with the hypothesis that the homoeobox sequence serves as a desig-

nator on nuclear destination is its location in the 3' exon of the *Drosophila* homoeo genes, which would give it an omni-present position in the variable gene products derived from the complex 5' splicing patterns of many of these genes. This hypothesis also offers a simpler explanation for the presence of multiple copies of similar sequences in the genomes of many species and obviates the necessity of invoking homologies in segmentation processes between organisms from diverse phyla. McGinnis *et al.*¹⁰, for example, have suggested that

If the conserved homoeo domain in fruit flies, frogs, mice, and humans is involved in segmental development, then it is possible that the segmentally organized animals in both the protostome and deuterostome classes had a common ancestor, and that the metameric body plan has evolved only once in evolution.

The phylogenetic soundness of this argument is open to question. Annelids and arthropods are fundamentally metameric animals in which the body is organized as a linear series of segments. In primitive forms, these segments and their repeated internal and external structures are very similar. In the evolution of insects these metameric segments have become highly differentiated from each other. Homoeotic genes can be viewed as having become progressively more elaborate as regulators of segment differentiation in this evolutionary lineage¹². While chordates do possess some serially repeated structures, most prominent of which are the somites, they are in no real sense metameric, as can clearly be seen from the most primitive group of chordates, the ascidians. While we agree that the protostomes and deuterostomes must share at some point a common ancestor, chordates are, in fact, not segmented animals and are fundamentally distinct from arthropods and other metameric animals both in body plan and in the developmental processes by which the body plan is achieved. The evolution of these developmental programs offers little evidence of meaningful homology¹².

We bring up this hypothesis for the sake of discussion; the answer of course will lie in data to come. Certainly a sequence of four to five basic amino acids is not a "universal" nuclear tag. Nuclear designator sequences for a number of other proteins have also been characterized; these comprise several different amino acid sequences placed in different places within the parent proteins¹³⁻¹⁵. In any event, if the homoeobox sequence indeed functions as a designator of nuclear destination, the possibility is testable by experiments such as those of Hall *et al.*¹⁵, who showed, by constructing fused gene products, that sequences from the nuclear protein yeast mating type factor Mat $\alpha 2$ are also able to confer nuclear localization on heterologous proteins. Finally, let us note the

reported homology between the homoeo domain protein sequence and both of the yeast mating type factors, Mat $\alpha 1$ and Mat $\alpha 2$ (refs 8, 9); certainly there is little argument for *Saccharomyces* as a segmented organism.

ELIZABETH C. RAFF
RUDOLF A. RAFF

*Institute for Molecular and Cell Biology
and Department of Biology,
Indiana University,
Bloomington, Indiana 47405, USA*

1. Bender, W. *et al. Science* **221**, 23-29 (1983).
2. Scott, M.P. *et al. Cell* **35** 763-776 (1983).
3. Kuroiwa, A., Hafen E. & Gehring, W.J. *Cell* **37**, 825-831 (1984).
4. Hafen, E., Kuroiwa, A. & Gehring, W.J. *Cell* **37**, 833-841 (1984).
5. Weiner, A.J., Scott, M.P. & Kaufman, T.C. *Cell* **37**, 843-851 (1984).
6. McGinnis, W., Garber, R.L., Wirz, J., Kuroiwa, A. & Gehring, W.J. *Cell* **37**, 403-408 (1984).
7. Carrasco, A.E., McGinnis, W., Gehring, W.J. & DeRobertis, E.M. *Cell* **37**, 409-414 (1984).
8. Loughton, A. & Scott, M.P. *Nature* **310**, 25-31 (1984).
9. Shepherd, J.C.W., McGinnis, W., Carrasco, A.E., DeRobertis, E.M. & Gehring, W.J. *Nature* **310**, 70-71 (1984).
10. McGinnis, W., Hart, C.P., Gehring, W.J. & Ruddle, R.M. *Cell* **38**, 675-680 (1984).
11. Kalderon, D., Richardson, W.D., Markham, A.F. & Smith, A.E. *Nature* **311**, 33-38 (1984).
12. Raff, R.A. & Kaufman, T.C. *Embryos, Genes, and Evolution* (Macmillan, New York, 1983).
13. Dingwell, C., Sharnick, S.V. & Laskey, R.A. *Cell* **30**, 449-458 (1982).
14. Dingwell, C. & Allen, J. *EMBO J.* **3**, 1933-1977 (1984).
15. Hall, M.N., Hereford, L. & Herskowitz, I. *Cell* **36**, 1057-1065 (1984).

SIR — The current research into DNA sequences involved in segment pattern formation in invertebrates¹⁻⁴ and their putative homologues in vertebrates⁵ raises a variety of questions concerning, first, the nature of the development process in vertebrates and invertebrates and, second, the nomenclature used to describe this process and its associated genomic sequences. Comparisons between invertebrate and vertebrate development may be, at this early stage, invidious, and we would like to take this opportunity to state the reason why we feel it to be dangerous to develop an all-encompassing paradigm for the problems of segmentation in higher organisms.

Much interest is concentrated on the homoeobox sequence common to several homoeotic genes in *Drosophila melanogaster*⁶. Similar sequences are found in vertebrates^{5,7,13}. The homoeobox sequence is also found in a *Drosophila* segmentation gene (*fushi tarazu*) which does not confer a homoeotic mutant phenotype^{6,8}. While the exact role of this sequence is unclear, in both invertebrates and vertebrates, it clearly shows a common association with the process of segmentation or metamorphism in arthropods.

Animals which are truly metameric or show metameric stages are confined to the invertebrates. Vertebrates show limited segmentation which is confined primarily to the muscles of the trunk and associated skeletal and nervous systems. Indeed, it is possible that vertebrate segmentation has developed independently of vertebrate groups in evolution^{9,10}. This raises the question of whether the homoeobox sequences, found in both invertebrates and vertebrates, plays, as has been suggested¹¹,