Virchows Archiv A Pathol Anat (1992) 420:291–294

*Virchows Archiv A*
*Pathological Anatomy and Histopathology*


*Editorial*

# What's in a homeobox

## The development of pattern during embryonic growth

**C.L. Berry**

Department of Morbid Anatomy, The Royal London Hospital, Whitechapel, London, E1 1BB, UK

Received January 20, 1992

No pathologist can have failed to notice the number of publications on the homeobox in recent years. These have provided tantalising glimpses of unsuspected interrelationships in growth and development, but are they a conceptual problem only or will they matter to practising pathologists? Our reaction to oncogenes may be a good topic for comparison; in the comparatively recent past they clearly seemed to some to be of greater importance as a concept in carcinogenesis than as a developing aspect of pathological practise, but the review by Wynford-Thomas (1991) indicates that this view can no longer be sustained. Homeoboxes appear to be an example of that most appealing type of concept, one which creates a unifying hypothesis with broad implications out of an area of confusion. But what are they? What do they do? How do they work? And what have they to do with growth and pattern formation in general? As homeobox activity appears to be a vital part of the developmental process it is worth putting the whole area of interest in its context.

First of all, what are they? Homeobox genes encode a highly conserved sequence of 60 amino acid residues, forming a domain in a set of proteins which bind to specific sites on the DNA of a target gene, thus regulating gene transcription. The polypeptide chain in the homeodomain consists of four helices, only one of which recognises a specific DNA sequence, but since this helix is almost identical in all homeodomain proteins, they all bind to similar DNA sequences. In man the genes are distributed in four different linkage groups on four different chromosomes; these groups are believed to have arisen during evolution as two duplications of chromosomal segments. Each of the complexes consists of a string of related genes all coding for transcription factors containing a homeobox sequence. The genes in the box are brought into action obeying the rule that the further the gene lies from the beginning of the complex in the chromosome the more posterior is its expression domain along the body axis (genes located at the 5'end of the complexes are expressed at the posterior end of the embryo and those toward the 3'end are expressed progressively more anteriorly). This pattern of arrangement of the genes in the chromosome in the same order in which they are expressed along the antero-posterior body axis is seen in the nematode *C. elegans* as well as in the insecta and mammalia. More than 40 homeobox genes have been described.

It appears that all mammals have at least four *HOM* gene complexes homologous to those of insects and two of these complexes, *Hox-2* and *Hox-5*, are known to be expressed in the central body axis of the mouse in the same way as in insects – further evidence that *Hox* expression provides cells with an indication of their cranio-caudal position (Graham et al. 1989). The same is clearly true in the proximo-distal configuration of the limbs where, in an extensive study, Dolle et al. (1989) mapped the expression of five members of the *Hox-5* gene complex by in situ hybridisation at different stages of development. Each gene had a specific spatial domain and time of activation within these domains and they nested like a set of Baboushka dolls in the order of the occurrence of the gene on the chromosome. Later in development, at the sites of potential joints the genes are down-regulated in a manner which suggests the activation of later acting gene products.

The initial discovery of the homeobox was made via the study of the mutations that affect segmentation in *Drosophilia*. Homeotic mutations alter the programme of a particular imaginal cell disc usually by causing it to follow the pathway of another. Thus a leg may develop where an eye should be; it will be a normal leg but in the wrong place. The development of the wings and haltaires is also affected by specific mutations; three of these (*bx, abx* and *pbx*) allow the development of an extra set of wings rather than haltaires. From the effects of these genes it is evident that their products act by restricting the developmental pathway of the discs and that their products may give "inappropriate" instructions, if altered by mutation.

Homeobox genes are frequently expressed in both induced and inducing cells. The formation of the brain and upper spinal cord is programmed by interactions

with the immediately contiguous mesoderm and the genes are clearly expressed in both neurectoderm and mesoderm. Their boundaries of expression are limited by rhombomeres (the segmented components of the hindbrain); for *Hox 2.9* the endoderm of the pharyngeal pouch will also express the gene. Indeed, each of the anterior branchial arches is populated by neural crest cells that express a unique set of homeobox genes determined by the area of the neural plate from which they are derived. Similar regionality is seen for the kidney and in the limbs (Chuong et al. 1990; Vogels et al. 1990).

From these data it appears that the homeobox determines cephalo-caudal gradients and proximo-distal gradients in different circumstances but that, in each case, a positional message is conveyed. The use of similar mechanisms across widely differing animal types suggests powerful reasons for conservation of an effective mechanism in development. Is this really what has happened? In order to determine this it is necessary to look closely at the way in which different classes use these genes and some general developmental points must be considered.

Embryos are not simply spatially arranged sets of differentiating cells. Morphogenetic movement and differentiation are separate, temporally related activities and developmental gene regulatory systems are necessary, firstly to organise the production of cell lineages (groups of cells with a defined mitotic history), then to position these with regard to the body plan and, finally, to carry out cell specification. Davidson (1991) has pointed out that important differences exist between the way in which embryos of different classes make use of highly conserved genes and that these differences may be of fundamental importance in morphogenesis. His arguments will be considered in detail below, but they provide a clear background against which pathologists can examine the issues.

The processes of organogenesis and morphogenesis are probably identical in all animal forms more complex than sponges. Homeobox activity appears to determine the antero-posterior determination of form in widely divergent animal forms by the regional expression of genes controlling locally inductive morphogenesis. However, close study of the way in which an egg is converted into an embryo reveals important differences in the way this happens. In nematodes, deuterostomes and neogastropods all initial cleavage is invariant and the lineages derived from specific blastomeres display predictable invariant fates – hence the many maps of cell predestination published for Amphiba and once common in embryology texts. But there are shades of grey. Some cell lineages are *autonomously* determined and will not vary whatever cell a blastomere finds itself next to, while others are *conditional;* the blastomere gives rise to different cells depending on local signalling. These embryos were classifed by Davidson as type I; a major characteristic is a direct relationship between the polarity of the egg and *an* axial feature of the body plan. Importantly, cell type specification precedes any large-scale embryonic cell migration.

Vertebrate and insect embryos differ from type I embryos. No vertebrate embryo has an invariate cell lineage and given structures are always composed of cells with differing lineages. Even after a significant number of cell divisions there remain cells which can give rise to a number of cell types (demonstrated for the zebrafish, chick and mouse). As an example, consider the neural crest, where the fate of individual cells is largely determined environmentally by influences acting during their migration. In this type of embryo, cell-type specification generally occurs after gastrulation and cell-type specific molecular markers of differentiation can only be detected at gastrulation or later (neurulation). There is a mechanism for regional identification, important since many of the cell types appear in all regions. This is mediated by diffusable intercellular morphogens such as FGF, TGF$\beta$ and retinoic acid. After the specification of regions such as the neural plate, head, and dorsal and ventral mesoderm the morphogenetic programmes begin to act. These embryos are in Davidson's type II.

The insects are properly in a separate type III. Initial development takes place in a syncytium. Cellular interaction cannot occur and after cellularisation the subsequent cell lineage is variable. The future metameric pattern of *Drosophila* is established before cellularisation is complete by an anteroposterior gradient of maternally derived gene product which is interpreted at the the level of regulatory DNA-protein interaction by zygotically expressed gap genes. These establish a broad series of overlapping domains and local concentrations are active in the syncytial blastoderm nuclei to produce transverse stripes of expression of pair rule genes. These control one another and segment polarity genes. Homeotic genes are initially controlled by gap gene products and later by segmental polarity gene products. As these events take place in a syncytium the concentration clines at the nuclei of diffusable macromolecular regulatory products and *not intercellular interactions* are critical. The mechanisms described for type I and II embryos are not available to insects but the end-point is identical – in Davidson's words "the generation of an assembly of spatially organised cells according to the body plan, that can now mount the intercellular functions required for cell type specification and morphogenesis."

In type 1 embryos regional control functions will only be used in postembryonic development – lineage, cleavage pattern and gastrulation determine early form. In type 11 embryos long-range diffusable morphogens apparently determine the broad organisation of body plan and morphogens such as retinoic acid exert their effects by homeobox gene activation – TGF-$\beta$ and bFGF are necessary for the initial axial specification of mesoderm with subsequent homeobox activity. In type 111 embryos a different process is followed, avoiding diffusable growth factors, short-range cell interactions or lineage-determined events. Homeobox genes generate the body plan ab initio rather than being used to continue morphogenesis after the establishment of this plan by other means; these include the control of expression of ligands and receptors.

Phylogenetically it seems probable that the original type of embryogenesis was type I. It seems that this mod-

el is easily modified as in the sea urchins, for example, direct development rather than development via an intermediary larval form has apparently evolved in six of the ten echinoid orders and may be found in the same genus. This type of change involves the use in embryonic development of some of the processes used in larval morphogenesis; there are changes in the timing and order of internal organ system morphogenesis and in the timing of expression of lineage-specific marker genes.

## The effects of altered homeobox expression

Ectopic expression or inhibition of expression of regional homeotic genes has profound morphogenetic effects. Wright et al. (1989) showed that the upper spinal cord could be compelled to develop as hind-brain if an antibody to part of *XlHbox-1* (*Xenopus*) was injected into a fertilised egg; a similar effect could be induced by over-expression of another part of the gene. The authors considered that the effect was produced by effects on the mesoderm underlying the CNS, but it is important to note that the disturbance of morphology produced was consistent with a general finding that the effect of ectopic homeobox expression is to transform the body segment into a more anterior type. That the genes act on regional organisation and not on differentiation is shown by the experiments of Harvey and Melton (1988), who found normal somitic muscle differentiation despite disruption of somite pattern formation in *Xenopus* after injecting synthetic transcripts of *Xhox-1A*. Over-expression of *Hox-1.4* in transgenic mice with elevated transcription levels of the gene in the mesenchymal layer of the embryonic gut resulted in a phenocopy of Hirschprung's disease with failure of the innervation of the large bowl, due either due to failure of neural crest cell migration or of appropriate development of gut mesenchyme (Wolgemuth et al. 1989). The DiGeorge syndrome is modelled by germline deletion of the *Hox-1.5* gene (Chisaka and Capecchi 1991).

## Oncogenes and growth factors in development: their relationship to homeoboxes

There appears to be a significant interaction of growth factors and proto-oncogenes in development. Many proto-oncogenes (*int-1, int-2,* c-*fms,* c-*myc,* N-*myc,* c-*jun* and *jun B*) are expressed in specific temporal and spatial patterns during development. Levels of *p53* appear to be regulated by differential gene expression as well at the post-transcriptional level and Schmid et al. (1991) have provided evidence for a particular role for this gene in differentiation; it has been shown that its level of expression in fetal mice is not well correlated with cell proliferation. However, up-regulation of *p53* gene expression may be necessary to inhibit cell cycle progression and allow terminal differentiation at many sites. This appears to be an exception; in general, oncogenes are probably important in ensuring the adequacy of anlage in terms of cell numbers and lineage. Do growth factors act in the same way? Some appear to be more important in the earliest stages of embryogenesis; others show hybrid functions. The *int-2* proto-oncogene family encodes a number of proteins related to FGF, a powerful inducer of mesoderm in the Amphibia, a fundamental step in embryogenesis. Far downstream, Represa et al. (1991) have shown that *int-2* oncoproteins are inducers of inner ear formation, a function which can be mimicked by basic FGF in the absence of *int-2* expressing rhombencephalon. Here both proteins apparently act by increasing the cell proliferation rate, a necessary step in the invagination and infolding which is such an important feature of development of the ear.

To consider a different group, the TGF-$\beta$s are a highly conserved family of proteins (60–80% homology of the C terminal end of the mature proteins). There are five vertebrate members, each belonging to larger superfamilies. They act to stimulate or suppress cell growth depending on cell type but have differing potencies; TGF-$\beta$2 is a powerful inducer of mesoderm in *Xenopus* but TGF-$\beta$1 can act in this way only in the presence of FGF. There is abundant evidence of differential expression in different tissues in murine embryogenesis (see Robinson et al. 1991, for bibliography). mRNAs for TGF-$\alpha$, TGF-$\beta$ and PDGF-A are present in pre-implantation mouse embryos and EGF and TGF-$\alpha$ stimulate the growth of two-cell embryos to the blastocyst stage; Dardik and Schultz (1991) have shown that TGF-$\alpha$ and EGF are vital in the expansion of the blastocoele cavity – a necessary event for morphogenesis.

## Pathological implications

The significance of this better understanding of basic processes is that we can at last begin to unravel the complex interactions between the genetic background of an individual and the environment – interactions which must, from epidemiological data on human malformation, be critical in dysmorphogenesis.

The major abnormalities of the central nervous and cardiovascular systems – which together form the majority of the burden of malformations in Man – are determined very early in development when the genes discussed above are active. It seems that rather than a simple programme of development fixed, as a blueprint, in the genes, what we have is the establishment of an embryo from the pre-embryonic form via the major inductive processes regulated by growth factors. Subsequently there is the establishment of axis and form where the homeoboxes are critical (in syncytia or multicellular frameworks), the development of anlage where cell proliferation rates are controlled by proto-oncogenes and finally signals for differentiation, which may come from any of these (and other) classes of gene. Mesodermal induction failures are probably of major significance in CNS anomalies (see Berry 1992) but morphological abnormalities due to failure of appropriate homeobox activity appear to be rare. The growth of anlage is vulnerable to environmental influences and interference with

differentiation is readily achieved with promoting agents such as the phorbol esters. Knowledge of the type of factor which is likely to be affected in a particular case may help in the identification of exogenous agents having adverse developmental effects.

## References


Berry CL (1992) A view of neurospinal dysraphism. Virchows Arch [A] (in press)

Chisaka O, Capecchi MR (1991) Regionally restricted developmental defects resulting from targeted disruption of the mouse homeobox gene *Hox-1.5*. Nature 350:473–479

Chuong C-M, Oliver G, Ting SA, Jegalian BG, Chen RM, De Robertis EM (1990) Gradients of homeoproteins in developing feather buds. Development 110:1021–1030

Dardik A, Schultz RM (1991) Blastocoele expansion in the preimplantation mouse embryo: stimulatory effect of TGF-alpha and EGF. Development 113:919–930

Davidson EH (1991) Spatial mechanisms of gene regulation in metazoan embryos. Development 113:1–26

Dolle P, Izpisua-Belmonte J-C, Falkenstein H, Renucci A, Duboule D (1989) Co-ordinate expression of the murine *Hox-5* complex homeobox-containing genes during limb pattern formation. Nature 342:767–772

Graham A, Papalopulu N, Krumlauf R (1989) The Murine and Drosophila homeobox genes have common features of organisation and expression. Cell 57:367–378

Harvey RP, Melton DA (1988) Microinjection of synthetic *xhox-1A* homeobox mRNA disrupts somite formation in developing *Xenopus* embryos. Cell 53:687–697

Represa J, Leon Y, Miner C, Giraldez F (1991) The *int-2* protooncogene is responsible for the induction of the inner ear. Nature 353:561–563

Robinson SD, Silberstein GB, Roberts AB, Flanders KC, Daniel CW (1991) Regulated expression and growth inhibitory effects of transforming groth factor-Beta isoforms during mouse mamary gland development. Development 113:867–878

Schmid P, Cox D, Bilbe G, Maier R, McMaster GK (1991) Differential expression of TGF $\beta$1, $\beta$2 and $\beta$3 genes during mouse embryogenesis. Development 111:117–130

Vogels R, De Graaff W, Deschamps J (1990) Expression of murine homeobox-containing gene *Hox-2.3* suggests multiple time-dependent and tissue-specific roles during development. Development 110:1159–1168

Wolgemuth DJ, Behringer RR, Mosteller MP, Brinster RL, Palmiter RD (1989) Transgenic mice over-expressing the mouse hoemobox containing gene *Hox-1.4* exhibit abnormal gut development. Nature 337:464–474

Wright CVE, Cho KWY, Hardwicke J, Collins RH, De Robertis EM (1989) Interference with function of a homeobox gene in *Xenopus* embryos produces malformations of the anterior spinal cord. Cell 58:81–93

Wynford-Thomas D (1991) Oncogenes and anti-oncogenes; the molecular basis of tumour behaviour. J Pathol 165:187–201