

REVIEWS

- J. Biol. Chem.* 264, 14312–14317
- 13 D'Mello, N.P. *et al.* (1994) *J. Biol. Chem.* 269, 15451–15459
- 14 Chen, J.B. *et al.* (1990) *Mol. Microbiol.* 4, 2081–2086
- 15 Sun, J. *et al.* (1994) *J. Biol. Chem.* 269, 18638–18645
- 16 Kataoka, T. *et al.* (1984) *Cell* 37, 437–445
- 17 Toda, T. *et al.* (1985) *Cell* 40, 27–36
- 18 Kennedy, B.K. *et al.* (1995) *Cell* 80, 485–496
- 19 Hartwell, L.H. (1980) *J. Cell Biol.* 85, 811–822
- 20 Ivy, J.M., Klar, A.J.S. and Hicks, J.B. (1986) *Mol. Cell. Biol.* 6, 688–702
- 21 Rine, J. and Herskowitz, I. (1987) *Genetics* 116, 9–22
- 22 Aparicio, O.M., Billingham, B.L. and Gottschling, D.E. (1991) *Cell* 66, 1279–1287
- 23 Rine, J. *et al.* (1979) *Genetics* 93, 877–901
- 24 Buck, S.W. and Shore, D. (1995) *Genes Dev.* 9, 370–384
- 25 Marcand, S. *et al.* (1996) *Genes Dev.* 10, 1297–1309
- 26 Morretti, P. *et al.* (1994) *Genes Dev.* 8, 2257–2269
- 27 Cockell, M. *et al.* (1995) *J. Cell Biol.* 129, 909–924
- 28 Shore, D. and Nasmyth, K. (1987) *Cell* 51, 721–732
- 29 Conrad, M.N. *et al.* (1990) *Cell* 63, 739–750
- 30 Lustig, A.J., Kurtz, S. and Shore, D. (1990) *Science* 250, 549–553
- 31 Olovnikov, A.M. (1973) *J. Theor. Biol.* 41, 181–190
- 32 Schwartz, H.S., Dahir, G.A. and Butler, M.G. (1993) *Cancer Genet. Cytogenet.* 71, 132–138
- 33 Vaziri, H. *et al.* (1993) *Am. J. Hum. Genet.* 52, 661–667
- 34 Harley, C.B., Futcher, A.B. and Greider, C.W. (1990) *Nature* 345, 458–460
- 35 Allsopp, R.C. *et al.* (1992) *Proc. Natl. Acad. Sci. U. S. A.* 89, 10114–10118
- 36 Feng, J. *et al.* (1995) *Science* 269, 1236–1241
- 37 Counter, C.M. *et al.* (1992) *EMBO J.* 11, 1921–1929
- 38 Palladino, F. *et al.* (1993) *Cell* 75, 543–555
- 39 D'Mello, N.P. and Jazwinski, S.M. (1991) *J. Bacteriol.* 173, 6709–6713
- 40 Linnemans, W.A., Boer, P. and Elbers, P.F. (1977) *J. Bacteriol.* 131, 638–644
- 41 Fields, S. and Herskowitz, I. (1987) *Mol. Cell. Biol.* 7, 3818–3821
- 42 Miller, A.M. (1984) *EMBO J.* 3, 1061–1065
- 43 Byers, B. and Goetsch, L. (1975) *J. Bacteriol.* 124, 511–523
- 44 Nasmyth, K. (1983) *Nature* 302, 670–676
- 45 McConnell, S.J. *et al.* (1990) *J. Cell Biol.* 111, 967–976

B.K. KENNEDY (bkennedy@mit.edu) AND **L. GUARENTE** (leng@mit.edu) ARE IN THE DEPARTMENT OF BIOLOGY, MIT, CAMBRIDGE, MA 02139, USA.

PERSPECTIVES

The central problem of the development of pattern and form is how genetic information can be translated in a reliable manner to give specific and complex multicellular organisms. Pattern formation is the process by which the spatial aspect of cellular differentiation is organized; how, for example, cartilage and muscle differentiate in just the right place during the development of the limb. Pattern formation, thus, focuses less on cell differentiation itself. After all, the difference between our arm and that of a bat, seal, or hippopotamus is not due to differences in muscle and cartilage differentiation but to their spatial pattern and later growth.

History

The idea that the behaviour of a cell is determined by its position in the embryo goes back to Hans Driesch at the end of the last century¹. Driesch discovered regulation in sea urchin embryos; that is, that the embryo could still develop into a normal larva even when parts were removed. Driesch, by ignoring inconvenient results, concluded that the embryo was an equipotential system – all cells were totipotent and there was a coordinate system that specified cell position and so determined its fate. He could not imagine that such a system could operate based on any known physical principles and invoked mystical forces.

A very important advance in understanding how patterns might develop came from one of my heroes, Thomas Hunt Morgan, at the beginning of the century. It was Morgan who first appreciated the possibility of gradients controlling pattern formation. This idea came from his studies on regeneration². He clearly stated how a gradient in some property could provide both polarity in a regenerating system and how responses at different

One hundred years of positional information

LEWIS WOLFERT

One mechanism by which spatial patterns of cell differentiation could be specified during embryonic development and regeneration is based on positional information. Cells acquire a positional value with respect to boundaries and then interpret this in terms of a programme determined by their genetic constitution and developmental history. The signals and the molecular basis of such a system have both been rather well conserved. Recent work has shown that cells can respond to quite small differences in the concentrations of molecules whose concentration could provide positional information.

thresholds could pattern the system. But it was as if his nerve failed him and he periodically abandoned these ideas in favour of an impenetrable mechanical model. However, ideas on gradients were taken up by C.M. Child (not one of my heroes), who found gradients of one sort or another in many systems and in metabolism in particular. Gradients were fashionable in the 1930s (Ref. 3) and the Swedish school of embryologists, which included Runnstrom and Horstadius, made great use of them to explain their results on early sea urchin development. But little attention was given to how the gradients could be set up, maintained, or how they could generate pattern.

By the 1960s, with the exception of the work on the insect epidermis by Stumpf, Locke, Lawrence and others (see Ref. 4), there had been little further progress

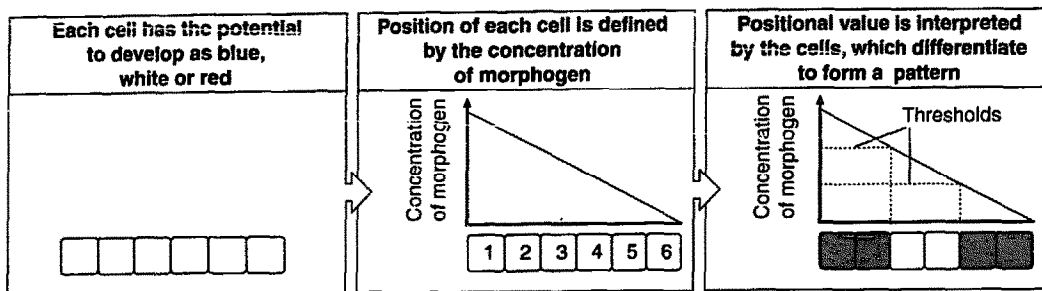


FIGURE 1. The French flag model for thinking about pattern formation. The cells in a line can differentiate as blue, white or red. They can be programmed to make a French flag pattern if they have their position specified and interpret this information. One mechanism can use a morphogen gradient and thresholds. The boundary values are crucial and the slope of the gradient gives the system its polarity. (Reproduced, with permission, from Ref. 26.)

in studying pattern formation. It was a rather neglected area. Gradients became out of fashion, and development was largely seen as a sequential process involving the synthesis of a large number of different proteins; all that was needed was to work these out in detail.

My own involvement with gradients came from experience with sea urchin morphogenesis⁵ and regeneration of hydra⁶. I was convinced that there must be general mechanisms for patterning during development and regeneration; that there was a general set of principles underlying the translation of genetic information into pattern and form. I was also sure that one first had to understand the processes at the cellular level before one could ask the right molecular questions. To focus my ideas, I formulated the French flag problem (Fig. 1). What mechanism would ensure that a line of totipotent cells, no matter how long, would always have a French flag pattern – a third blue, a third white and a third red? Positional information provided a general solution⁷.

The basic idea of positional information is that there is a cell parameter (positional value) that is related to a cell's position in the developing system⁸. It requires a coordinate system with respect to which the cells have their position specified. The cells then interpret their positional value by differentiating in a particular way. This differentiation may involve developing as a particular cell type or state, or it might involve changes in growth or motility.

For a one-dimensional system like the French flag, all the necessary features can be provided by a monotonic decrease in the concentration of a chemical – a morphogen (Fig. 1). The concentration of the morphogen at any point then provides a measure of distance from the boundary, and the slope of the concentration gradient effectively provides the polarity. Positional information can also be established by other mechanisms, for example by measuring the time spent in a region, such as the progress zone in the developing limb⁹. It is not necessary that positional information be specified by a gradient; the essential feature is that there is some cell parameter that reflects distance from a boundary. Position can be specified by the sequential activation of a string of genes. It is also interesting that most patterning occurs in two-dimensional sheets of cells and different mechanisms are probably used for specifying position along the two axes.

Among the potentially attractive features of positional information was that it could provide a unifying concept for understanding the development and regulation of various patterns. The only cell-to-cell interactions that are, in principle, required are those necessary to specify position. Furthermore, the same signals and positional values may be used to specify different patterns, the differences arising from developmental history and from genetic constitution.

What should not be regarded as a system of positional information? If each cell in a developing system has a unique specification, this does not necessarily mean that these specified states arose through positional information in the sense of a coordinate system. Thus, in the development of the nematode *Caenorhabditis elegans*, most cells have a unique specification and position by virtue of their lineage, but there are no boundaries and no measurement of position with respect to them. Again, the development of the eight photoreceptors in the ommatidia of the *Drosophila* eye depends on sequential induction⁴. The cells' positions with respect to their neighbours is somewhat like that in folk dancing or a rugby scrum rather than in a coordinate system.

When I looked at the sizes of what were probably positional fields, I was surprised to find that all positional fields are small, none being more than about 0.5 mm in maximum linear dimension. For example, hydra is less than 60 cells long, the sea-urchin gastrula about 30, and the early chick limb bud less than 100. The other characteristic feature is that the times required to specify position appear to be on the order of hours. It was these two features that lead Crick¹⁰ to propose the existence of a diffusible morphogen for setting up positional fields. One of the effects of thinking in terms of positional information was that one was driven to become more mechanistic and quantitative.

The major advances in our understanding of positional information over the last 30 years have come from molecular studies that have identified positional signals and the molecular basis of positional values, and have demonstrated that cells can respond to threshold concentrations.

Positional fields

By far the clearest demonstration of positional signals in a developing system – clear in the sense that

PERSPECTIVES

they can be directly visualized rather than being inferred from other properties – is in the *Drosophila* embryo⁴. *Drosophila*, however, is a bit peculiar because in the early embryo there are no cell walls and the nuclei share a common cytoplasm. Along the anteroposterior axis, Nusslein-Volhard and Driever showed that a gradient in the protein bicoid is set up and this provides positional information for early patterning. The gene *hunchback* is activated above a threshold concentration of bicoid and the gradient of hunchback provides further positional signals for other genes. Along the dorsoventral axis, the protein product of the *dorsal* gene becomes graded in nuclei with the high point ventrally. Again, at particular thresholds, other genes are activated; for example *twist* and *snail* are activated at high nuclear concentrations of dorsal and they go on to specify the mesoderm.

A feature of some positional fields is that the same pattern develops even when the field varies considerably in size. The classical case is Driesch's demonstration that each of the two cells that result from the first cleavage in the sea urchin will develop into small but normal larvae. Twinning is quite common in mammals, and in humans splitting can occur even when there are many hundreds of cells in the embryo. The patterning of such regulative fields requires that both ends should act as boundaries and provide signals. There are signalling boundary regions at both ends of the anteroposterior axis of the *Drosophila* embryo and also two at opposite ends of the dorsoventral axis. Further good examples are the patterning of the early *Xenopus* embryo¹¹ and the vertebrate neural tube where there are ventral and dorsal signals¹². Also, the classical work on sea urchins was based on two gradients, one animal, the other vegetal.

There is very good evidence for positional signals in the development of the vertebrate limb¹³ and the insect wing¹⁴. Moreover, they share some remarkable similarities (Fig. 2). Patterning along the anteroposterior axis of the chick limb bud involves a positional signal from the polarizing region at the posterior margin. This signal can specify the pattern of cartilage and muscle differentiation. For example, when the polarizing region (P) is grafted to the anterior margin, a complete set of additional digits can be specified to give (P)432234, while when placed in the middle, the pattern is 234(P)434. Sonic hedgehog is an excellent candidate for such a signal: it is expressed in the polarizing region and increasing concentrations applied locally to the anterior margin result successively in the specification of digits 2, 3 and 4 (A. McMahon and C. Tickle, pers. commun.). Such experiments can be understood in terms of a model in which the signal provides a concentration gradient of a morphogen¹⁵.

In the development of the insect wing, patterning along the anteroposterior axis results from signals at the

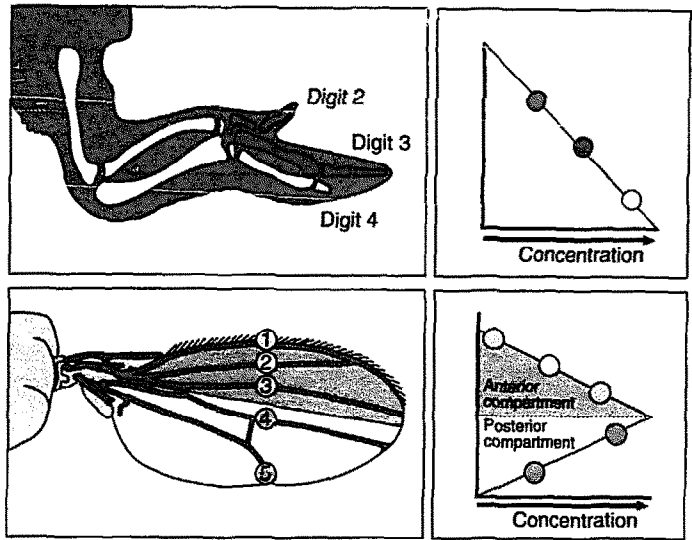


FIGURE 2. The chick wing and the insect wing both seem to use a positional signal – a morphogen to pattern the structure along the anteroposterior axis. In the chick the signal comes from the polarizing region. The signal in the insect is from the compartment boundary and is interpreted differently in the anterior and posterior compartments. (Reproduced, with permission, from Ref. 26.)

boundary that divide the wing imaginal disc into anterior and posterior compartments, as predicted by Meinhardt¹⁵. Along this boundary the gene *decapentaplegic* (*dpp*) is expressed. The formation of a boundary expressing *dpp* is dependent on *hedgehog*, which is expressed throughout the posterior compartment interacting, at the boundary, with cells in the anterior compartment. The veins – like the digits of the vertebrate limb – provide good markers for anteroposterior patterning: vein 1 being most anterior and vein 5 most posterior. The pattern of the veins is under the control of DPP protein in a concentration-dependent manner^{16,17}. The results of ectopic expression of *dpp* can be interpreted in terms of a symmetrical DPP gradient with its high point at the compartment boundary. While the concentration profile will be similar in anterior and posterior compartments, they will be interpreted differently. So, when a new source of DPP (D) is expressed in an anterior position, the pattern of the veins could now be 123(D)345.

Another system where these same positional signals have been shown to pattern the structure is the vertebrate neural tube¹². On the ventral side, Sonic hedgehog most likely acts as a positional signal in a concentration-dependent manner, while on the dorsal side, relatives of DPP (members of the TGF- β family) provide the signal.

One of the attractive features of positional information is that it suggests the possibility of universality. That is, in principle, the same coordinate system and the same signals could be used again and again, even in the same embryo. As just pointed out, the anterior and posterior compartments of the insect wing may interpret the same signal in different ways. Again, the signals in the forelimb and hindlimb of vertebrates are the

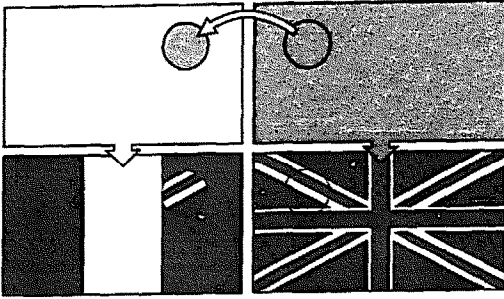


FIGURE 3. Cells can interpret their positional values according to their genetic constitution. So, if a small group of cells from a Union Jack are placed in a developing French flag they will make a Union Jack pattern corresponding to their new position. Similar results have been obtained with the leg and antenna of *Drosophila*. (Reproduced, with permission, from Ref. 26.)

same, as they are in the developing and regenerating amphibian limb¹⁸.

Thresholds

One of the major difficulties that was consistently raised against gradients underlying pattern formation was how they could be interpreted. The interpretation of a gradient so as to give discrete regions required the cells to accurately and reliably recognize particular concentrations of some morphogen. For the great embryologist Hans Spemann¹⁹ this seemed to pose an almost insuperable problem.

There is now excellent evidence that cells are very good at recognizing thresholds. A good example comes from studies on the development of the mesoderm in *Xenopus*. One of the factors that can induce mesoderm in presumptive ectoderm is activin. Treating presumptive ectoderm with increasing concentrations of activin results in a succession of mesodermal tissues being induced^{20,21}. At the lowest concentration, haemopoietic tissue is induced, while at the highest concentrations notochord is induced. Small increases in activin concentration can change the cell's response. Similarly, the concentration of DPP in the insect imaginal disc can activate specific genes at threshold concentrations^{16,17}.

Positional values and differentiation

A central feature of positional information is that the positional values of the cells are independent of how the cells will differentiate. Differentiation and positional value are dissociated. In principle, in insects, any epidermal structure, such as a sensory bristle, could develop at any position in the epidermis. There is, thus, no absolute relationship between the observed pattern and the underlying positional values; the same set of positional values can specify many different patterns.

The unconstrained relationship between positional value and differentiation is fundamental to the concept of positional information and is seen particularly clearly in genetic mosaics in insects⁴. Numerous other studies also show that the cells interpret their position according to their genotype and developmental history. It is as if cells programmed to develop into the Union Jack were placed in a developing French flag – they still develop into a part of the Union Jack but according to their new position (Fig. 3). A good example is the development of mesoderm taken from the presumptive thigh region of the chick leg bud and grafted to the tip of the wing bud. The tissue acquires a more distal positional value but interprets this with respect to the leg programme and so toes develop. Such mosaics also illustrate the identity of the positional values in different parts of the embryo. The same is true for the imaginal disks of insects; for example, the positional values are the same in the antenna and the leg.

In systems that use positional information, that information has to be recorded and interpreted. A major discovery has been the *Hox* gene complexes. These are prime candidates for recording positional value in all animals²². In *Drosophila*, along the anteroposterior axis, expression of the *Hox* gene complex provides each segment with a unique identity. Related genes in many other animals have been identified and they, too, seem to provide regions along the anteroposterior axis with a positional value (Fig. 4). Consistent with the idea that *Hox* genes provide positional values, mutations in these genes can cause localized abnormalities or even homeotic transformations, where one region (e.g. a vertebra) acquires the identity of a neighbouring region. While many *Hox* genes are also expressed in the developing limb, their role is much less easily understood. Induced mutations and knockouts can have varying

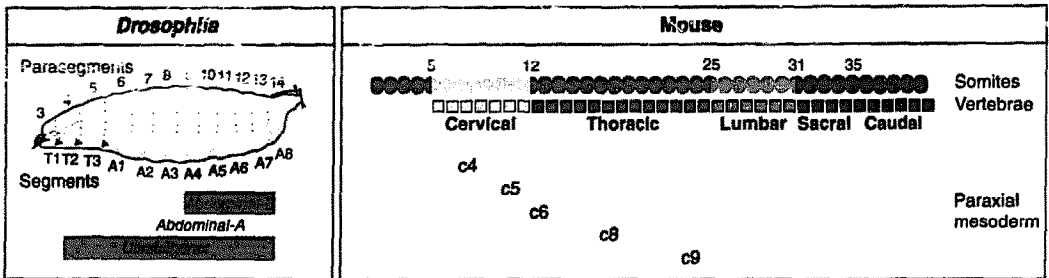


FIGURE 4. *Hox* genes provide positional values. In *Drosophila*, the three genes shown (*Abdominal-A* and *-B*, and *Ultrabithorax*) specify thoracic and abdominal segments. In the mouse, four gene complexes provide positional values, but only the *Hoxc* complex is shown here²⁷. Together, the four gene complexes specify positional values and so the pattern of development of the different vertebrae. The anterior borders of gene expression are sharp, but the posterior borders are much less well defined and extend some way to the posterior end. (Reproduced, with permission, from Ref. 26.)

PERSPECTIVES

effects suggesting a quite complex dynamic relation between *Hox* gene expression and the limb pattern.

Regeneration and intercalation

Regeneration can also be viewed in terms of positional value. For example, morphallaxis (regeneration without growth) and epimorphosis (regeneration linked with growth) can be formally distinguished within this framework (Fig. 5). Regeneration by morphallaxis would first require specifying new boundary regions at the cut edges and new positional values would be specified with respect to the new boundaries. With epimorphosis, new positional values would be linked to growth from the cut surface.

Studies on regeneration also provide good evidence that cells have positional values that can be changed. A wide variety of experiments strongly suggest that in, for example, the limbs of insects and vertebrates that can regenerate, there are sets of positional values along the main axis as well as circumferentially²³. Intercalation of intermediate values occurs when noncontiguous regions are placed next to each other to smooth out those disparities.

That cells have positional values unrelated to their differentiated state is again shown by regeneration of vertebrate limbs and of hydra. Not only do urodele vertebrate limbs show intercalation of the type just described, but regeneration can best be understood in terms of generating a new set of positional values continuous with the level of the cut. Retinoic acid can cause cells to alter their positional values so that they acquire more proximal values²⁴. So, in an animal treated with retinoic acid, regeneration from the level of the wrist can give rise to a whole new limb. Moreover, single marked cells that have acquired a more proximal positional value actually end up in more proximal regions. In hydra, too, regeneration is best understood in terms of establishing positional values corresponding to boundaries at the cut surface⁶.

Prospects

There are many details relating to pattern formation involving positional information that remain to be worked out. The most notable relates to interpretation. We still have a rather poor appreciation of the processes whereby positional values as specified by *Hox* genes result in the emergence of pattern. This is particularly striking in the vertebrate limb: what are the downstream targets of the genes, and which cellular activities are being controlled? For example, with the cartilaginous elements is there a fundamental prepattern specified by a quite different mechanism that is being modified by *Hox* genes⁸? Is cartilage differentiation affected or only growth? Similar problems also exist with respect to imaginal disc development in insects. How does a mutation in just one gene alter the development of an antenna into a leg?

In spite of an enormous amount of work and persuasive evidence for signals acting over 10–20 cell diameters, the details of this process remain unknown. It is not easy to believe that if morphogen gradients do indeed control pattern formation, they should rely on simple diffusion through the extracellular space. Such a mechanism seems just too unreliable and a more sophisticated mechanism would surely have been 'invented'.

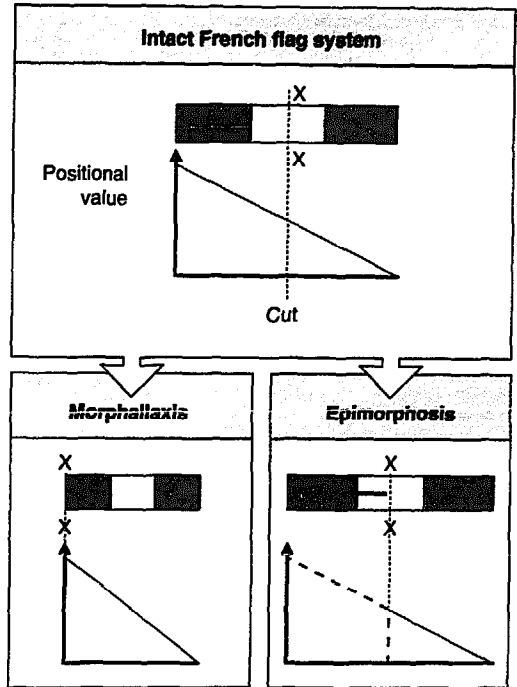


FIGURE 5. Regeneration by morphallaxis and epimorphosis. Morphallaxis involves the specification of a new boundary at the cut end (X) and then the alteration of intermediate positional values. With epimorphosis, there is the generation of new positional values from the cut end linked to growth. (Reproduced, with permission, from Ref. 26.)

Finally, we need to understand how developmental processes evolved²⁵. There is increasing evidence that evolution has been rather lazy. It seems as if once the basic mechanism for patterning a small region, such as a segment or an appendage, was found the main components, or at least principles, were used, with minor variations, again and again. But how primitive were gradients and positional values? And what was their origin and their relationship to mechanisms based on asymmetric cell divisions, and patterning, as in the nematode, which seems to develop on a cell-by-cell basis?

Acknowledgements

I am most grateful for the help and support of Cheryl Tickle and Jim Smith. I also remain indebted to the Nuffield Foundation for making possible my transfer from engineering to cell biology in 1955.

References

- 1 Wolpert, L. (1981) in *A History of Embryology* (Horder, T.J. et al., eds), pp. 347–361, Cambridge University Press
- 2 Wolpert, L. (1991) in *A History of Regeneration Research* (Dinsmore, C., ed.), pp. 201–217, Cambridge University Press
- 3 Huxley, J.S. and deBeer, G.R. (1934) *The Elements of Experimental Embryology*, Cambridge University Press
- 4 Lawrence, P.A. (1992) *The Making of a Fly*, Blackwell
- 5 Gustafson, T. and Wolpert, L. (1967) *Biol. Rev.* 42, 442–498

- 6 Wolpert, L., Hornbruch, A. and Clarke, M.R.B. (1974) *Am. Zool.* 14, 647-663
- 7 Wolpert, L. (1969) *J. Theor. Biol.* 25, 1-47
- 8 Wolpert, L. (1989) *Development Suppl.* 3-12
- 9 Sumnerbell, D., Lewis, J. and Wolpert, L. (1973) *Nature* 244, 492-496
- 10 Crick, F.H.C. (1970) *Nature* 225, 420-422
- 11 Cunliffe, V. and Smith, J.C. (1994) *EMBO J.* 13, 349-359
- 12 Placzek, M. and Furley, A. (1996) *Curr. Biol.* 6, 526-529
- 13 Cohn, M. and Tickle, C.A. *Trends Genet.* (in press)
- 14 Brook, N., Diaz-Benjumea, F. and Cohen, S. *Annu. Rev. Cell Biol.* (in press)
- 15 Meinhardt, H. (1983) *Dev. Biol.* 96, 375-385
- 16 Nellen, D., Burke, R., Struhl, G. and Basler, K. (1996) *Cell* 85, 357-368
- 17 Lecuit, T. *et al.* (1996) *Nature* 381, 387-393
- 18 Muneoka, K. and Sassoon, D. (1992) *Dev. Biol.* 152, 47-49
- 19 Spemann, H. (1938) *Embryonic Development and Induction*, Yale University Press
- 20 Green, J.B.A., New, H.V. and Smith, J.C. (1992) *Cell* 71, 731-739
- 21 Gurdon, J.B., Mitchell, A. and Mahony, D. (1995) *Nature* 376, 520-521
- 22 Slack, J.M.W., Holland, P.W.H. and Graham, C.F. (1993) *Nature* 361, 490-492
- 23 Bryant, S.V., French, V. and Bryant, P.J. (1981) *Science* 212, 993-1002
- 24 Pecorino, L.T., Entwistle, A. and Brockes, J.P. (1996) *Curr. Biol.* 6, 563-569
- 25 Wolpert, L. (1994) *Development (Suppl.)*, 79-84
- 26 Wolpert, L. *et al.* *Principles of Development*, Current Biology (in press)
- 27 Burke, A.C., Nelson, C.E., Morgan, B.A. and Tabin, C. (1995) *Development* 121, 333-346

L. WOLPERT (lwolpert@ucl.ac.uk) IS IN THE DEPARTMENT OF ANATOMY AND DEVELOPMENTAL BIOLOGY, UNIVERSITY COLLEGE, MEDAWAR BUILDING, GOWER ST, LONDON, UK WC1E 6BT.

Gene families arose during evolutionary history through a process of duplication of the ancestral gene followed by functional and structural specialization (divergence) of both copies¹⁻⁵. Gene families are widely accepted as a basis for classifying proteins, but their broader biological significance is less clear. Indeed, the existence of evolutionary homology between genes having dramatically different functional specificities is often dismissed as a curious but somewhat mysterious fact of life. Our knowledge of molecular genetics has now broadened so that it may be appropriate to re-examine this question and attempt to reach some general conclusions about the evolution of multigene families.

Gene duplication

Tandem genetic duplications in bacteria and bacteriophages occur spontaneously at a frequency of 10⁻³ to 10⁻⁵ per locus per generation, and can be of unlimited size⁶. Studies of the genetic basis of insecticide resistance suggest that spontaneous gene duplications occur in insects at a similar rate^{7,8}. Tandem duplications also occur frequently in mammalian cells, and can be reproducibly recovered by selection for resistance to cytotoxic drugs⁹, as well as during disease progression in untreated human tumors¹⁰. In general, removal of selection results in the rapid loss of the duplicated genes⁶⁻⁹. Spontaneous duplications can also occur by nontandem mechanisms, including chromosomal translocation, nondisjunction, polyploidization and retrotransposition^{1,5}. However, gene duplication and divergence within a gene family typically occur at a much slower rate (of the order of 10-100 million years or more per gene divergence event in any particular lineage of a gene family)^{5,11}. Thus, spontaneous gene duplications occur much too frequently to be the rate-limiting step in the evolutionary process of gene duplication and divergence. The rate-limiting step probably occurs at the level of natural selection (or genetic drift)

The coevolution of gene family trees

KARL J. FRYXELL

Gene duplication mutants arise spontaneously at a high rate in bacteria, bacteriophages, insects and mammalian cells, and are generally viable. Thus, the rate-limiting step in the evolutionary process of gene duplication and divergence was probably not gene duplication per se. Rather, it is likely that only a small fraction of all duplicated genes were retained, and were able to diverge into new specificities. Furthermore, gene duplications and functionally related gene families often show similarities in divergence dates, functional specificities, and phylogenetic tree topologies. These correlations suggest that the family trees of functionally related gene families coevolved because functionally complementary gene duplication and divergence events tended to be retained by natural selection.

after duplication. That is, the duplicated gene is likely to be lost unless it acquires a novel and useful function.

A hypothesis

The acquisition of a novel function by a duplicated gene could be facilitated by pre-existing heterogeneity in proteins that interact directly with the product of the duplicated gene. Thus, the successful duplication and divergence of one gene would provide an altered selective environment, which could facilitate the retention and divergence of duplicated copies of functionally interacting genes. The interacting genes could be duplicated through a tandem or nontandem mechanism, and could originate simultaneously, sequentially, or millions of years later. Regardless, the duplicated genes would tend to be lost (or mutated into pseudogenes) unless stabilized by natural selection, which would require