Genes That Guide Fruit Fly Development

Studies of genes that control fruit fly segmentation reveal a common DNA sequence that may provide clues to the regulation of early development

Fruit flies that have legs instead of antennae on their heads or extra pairs of wings may look as if they belong in a circus sideshow but they are more than mere genetic freaks. They are proving to be valuable guides to the mysteries of developmental regulation.

How a fertilized egg, a single cell, develops into a complex individual composed of many types of cells with differing compositions and functions has long been a puzzle. Investigators generally expect the solution to reside in the genes and have been looking for mutations that disrupt normal developmental patterns to lead them to the controlling genes.

And that is where the mutant flies come in. Their mutations are of the type called "homeotic," which cause cells to switch from one developmental fate to another. This can produce visible changes in the fruit fly body plan. One body segment, instead of developing its normal structural features, may acquire structures characteristic of another segment. Antennae may be replaced by legs, for example. The homeotic genes of the fruit fly are among the few that have been identified in any species that control events occurring in early development.

Within the past few years molecular biologists have been joining forces with more classic geneticists to probe the natures and activities of homeotic genes. In the fruit fly *Drosophila melanogaster*, many of these occur in two large gene clusters, the bithorax and Antennapedia (antenna-foot) complexes, both of which are located on the right arm of chromosome 3. The bithorax complex controls the development of the thoracic and abdominal segments while the Antennapedia complex primarily affects the thoracic and head segments.

Investigators first got a toehold in the bithorax complex, a major portion of which—nearly 200 kilobases of DNA—has now been cloned and mapped at the molecular level (*Science*, 25 September 1981, p. 1485; 1 July 1983, p. 23). More recently, two groups of investigators have cloned most or all of the Antennapedia complex, up to 300 kilobases of DNA in this case.* One of the major findings of the work is the identification of a short segment of DNA, about 180 base pairs in length, that is a component of several homeotic genes both in the Antennapedia and bithorax complexes.

The finding supports the hypothesis that the two complexes may have evolved from the same common ancestral gene. "The nicest thing is that the work confirms an evolutionary progression from one or a few genes that control segmentation," comments Edward Lewis of the California Institute of Technology. In addition, it provides a clue to how homeotic genes might act to control development and also a possible link by which the activities of the two complexes might be coordinated. Finally, very similar sequences have been found in species as diverse as the frog, the mouse, and the human and may serve as a means for identifying genes that help to regulate early development in these species.

> "The nicest thing is that the work confirms an evolutionary progression from one or a few genes that control development."

Cloning of the Antennapedia complex DNA was carried out primarily in the laboratories of Thomas Kaufman at Indiana University and Walter Gehring at the University of Basel. Both groups used methods originally developed for the bithorax complex by Welcome Bender of Harvard University and David Hogness of Stanford University School of Medicine. The investigators essentially "walk" along the chromosomal region of interest by collecting overlapping fragments of DNA.

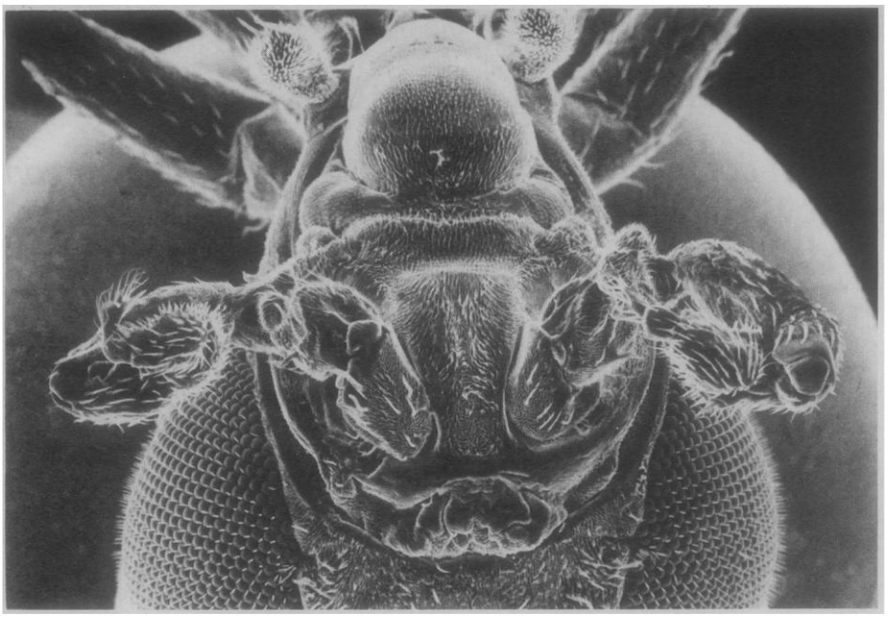
The Antennapedia complex contains at least three homeotic genes plus about five additional genes that are involved in development. One particular homeotic gene called *Antp* (for *Antennapedia*) turned out to be especially large. Kaufman and his colleagues found that *Antp* mutations that can cause the conversion of the antennae of adult flies into legs are distributed over some 100 kilobases of DNA. Moreover, both the Kaufman and Gehring groups showed that although the messenger RNA (mRNA) transcripts of the gene are only a few kilobases in length, they are derived from coding segments that are also widely spaced across more than 100 kilobases. Another gene that rivals *Antp* in size is the *Ultrabithorax* (*Ubx*) gene of the bithorax complex, which is about 73 kilobases long.

While mapping the *Antp* locus, the Kaufman and Gehring groups noted that one of their cloned probes also detected another gene in the Antennapedia complex, which is called *fushi tarazu* (Japanese for "not enough segments," abbreviated *ftz*). Strictly speaking, *ftz* is not a homeotic gene, although it is necessary for the correct development of segmentation in the fruit fly. Mutations in *ftz* are lethal when they occur in both copies of the gene, but the larvae partially develop before they die and can be seen to have only half the normal number of segments. "The *ftz* gene seems to be involved in counting the segments and the homeotic genes in determining the nature of the segments," notes Matthew Scott, who has recently moved from Kaufman's laboratory to the University of Colorado in Boulder.

For the *Antp* clone to recognize and bind to the *ftz* gene, the two DNA's must contain segments with similar nucleotide sequences. Gehring and his colleagues and Scott and Amy Weiner of the Indiana group have found the regions of similarity to consist of 180 base pairs of DNA near the 3′ ends of the genes. The *Drosophila* genome carries the sequence in five or more additional sites, including at least two loci in the bithorax complex (*Ubx* and *infraabdominal-2*) and another locus (*Deformed*) in the Antennapedia complex. Gehring has named the sequence the "homeo box" in view of its presence in many homeotic genes. It does not appear to be present in all of them, however. The Gehring group has not detected it in at least one gene of the bithorax complex.

Determination of the nucleotide sequences of homeo boxes from the differ-

*M. P. Scott, A. J. Weiner, T. I. Hazelrigg, B. A. Polisky, V. Pirotta, F. Scalenghe, T. C. Kaufman, *Cell* **35**, 763 (1983); R. L. Garber, A. Kuroiwa, W. J. Gehring, *EMBO J.* **2**, 2027 (1983); W. McGinnis, M. S. Levine, E. Hafen, A. Kuroiwa, W. J. Gehring, *Nature (London)* **308**, 428 (1984).



Antennapedia mutant

Scanning electron micrograph of a fruit fly with leglike structures growing from the head in place of antennae. [Courtesy of Rudolph Turner, University of Indiana]

ent genes shows that about 75 percent of the nucleotides are the same in all of them. "We find a fantastic degree of homology," Gehring says. "It is one of the most highly conserved sequences known."

The presence of the homeo box in the *ftz* gene and in homeotic genes from both the bithorax and Antennapedia complexes indicates that they may all be derived from a common ancestral gene. Lewis had previously proposed that the bithorax complex genes may have evolved by duplication of an ancestral gene, with the duplicated copies diversifying to specify the formation of different segmented structures. And Kaufman suggested that the Antennapedia complex is an anterior homolog of the bithorax complex.

Gehring and William McGinnis of the University of Basel have identified the homeo box in a number of species in addition to the fruit fly. These include the frog *Xenopus laevis*, the mouse, and the human. The sequence was not detected in nematodes, yeast, or sea urchins—all organisms that are not segmented.

The Gehring group, with Eddy de Robertis, also of Basel, has determined the nucleotide sequence of a frog homeo box and found that some 80 to 90 percent of the amino acids it encodes are identical to those encoded by the *Antp*, *ftz*, and *Ubx* homeo boxes.

Such close conservation of structure throughout evolutionary history suggests that the gene segment plays an essential role of some kind. There are indications that homeotic genes act by controlling the activities of other genes during development. One way in which they might do this is by directing the synthesis of proteins that bind to those genes and help to regulate their expression by turning them on or off.

DNA-binding proteins such as the histones, which are thought to be involved in the regulation of gene expression, are highly basic, a property also exhibited by the peptides encoded by the homeo boxes. Lysine and arginine residues constitute nearly one-third of the amino acids in the homeo sequences. Moreover, the structures bear a weak resemblance to those of the proteins encoded by the mating-type locus of yeast, which have been implicated in gene control in that species. "Some of the proteins encoded by homeotic genes may have a DNA-binding domain," Scott notes, "although known proteins that bind to specific DNA sequences are not especially basic in their DNA-binding domains." More work will be needed to confirm the hypothesis that the genes that contain the homeo box help to regulate developmental events by producing DNA-binding proteins.

In any event, there is evidence that the *ftz* gene is turned on very early in the development of the fruit fly embryo. During the first several rounds of mitosis, the nuclei of the fertilized egg divide synchronously within the inner yolk of the egg. Individual cells do not form at this stage. After seven rounds of division the nuclei begin migrating to the outer cortex of the egg where they form a monolayer. The nuclei undergo a few more rounds of division, for a total of 13 in all, and then become enclosed in membranes, giving rise to a single layer of cells (called the cellular blastoderm) that will form the various tissues of the fruit fly.

Developmental biologists have known for many years that the nuclei, before the migration, are still equivalent and totipotent; no matter where an individual nucleus ends up, it can develop normally for that region of the egg. However, the cells, by the time they form, are committed to contribute to specific segments of the animal.

The Gehring group first detected RNA transcripts of the *ftz* gene in the fruit fly embryo before the cells are formed, at about the 11th round of nuclear division. Weiner, Scott, and Kaufman also find that the transcripts are present very early in development, reaching a maximum concentration at the cellular blastoderm stage.

The transcripts are not evenly distributed throughout the length of the egg, according to Gehring and his colleagues, but are located in the region between about 15 and 65 percent of the length of the egg (measuring from the posterior end). By the 13th round of division, a segmented pattern can be seen in which bands of transcripts alternate with bands in which few transcripts are found, even though no obvious structural signs of segmentation are visible in the larvae.

As already mentioned, larvae homozygous for *ftz* mutations have half the usual number of segments. In normal larvae the *ftz* transcripts appear to be concentrated in precisely those embryonic regions that would give rise to the lost segments. "We see an arrangement of seven belts of hybridization that correspond perfectly to the missing segments," Gehring says. Continuous expression of the *ftz* gene is not necessary to establish the fruit fly segmentation pattern, however. By the time structural segmentation becomes visible, *ftz* transcripts are no longer detectable.

The results indicate that the *ftz* gene works very early in fruit fly development—it may be among the first genes of the embryo to be turned on—to help lay down the organism's segmentation pattern. "I think that the gene is a sensory element in the nuclei that tells them where they are in the embryo," Gehring explains.

Development may be controlled by a hierarchy of genes, Gehring postulates, with *ftz* occupying a high position in that hierarchy. Commitment of the blastoderm cells to develop into particular embryonic segments appears to depend on an interaction between the nuclei and the cortical cytoplasm in which they find

themselves after migrating. The positional information in the cytoplasm may in turn depend on the products of "maternal effect" genes, maternal genes that are expressed in the egg and have among their effects the establishment of the egg's anterior-to-posterior and dorsal-to-ventral axes.

Once *ftz* and other genes that are needed to establish the basic segmentation pattern are turned on, then the homeotic genes of the Antennapedia and bithorax complexes may be activated to direct those segments to diversify along appropriate paths. Activation of the genes that finally produce the different structural features of the fruit fly would be a fairly late event, possibly the result of the activities of the homeotic gene products, according to this view.

Whether or not the homeo box functions in higher species such as the frog, mouse, and human in the way postulated for it in the fruit fly is currently unknown. However, there are indications that genes bearing the boxes may function early in development in these species. The Basel workers have shown that such a gene in the frog is transcribed into mRNA by the late gastrula stage. The presence of homeo boxes may provide a handle by which some early developmental genes can be identified and studied, which would be a big help to researchers who wish to unravel the mysteries of development.—**JEAN L. MARX**

Computer Vision

This may be as close as AI has yet come to being a true science; but even so, no one really knows what it means to "see"

Legend has it that a certain pioneer in artificial intelligence research (AI) once gave a graduate student a little project for the summer: solve vision.

That was two decades ago.

One wonders if the student had a very good time that summer. Not only is his little problem of vision still unsolved, it is still one of the greatest challenges in AI. Vision systems do exist for industrial robots, for example, yet even now they tend to be primitive silhouette matchers with limited utility. And when the Defense Advanced Research Projects Agency (DARPA) recently launched its "Strategic Computing" initiative (*Science*, 16 December 1983, p. 1213), it estimated that another 10 years of concentrated effort would be required before an autonomous reconnaissance vehicle could "see" well enough to rove over unknown terrain.

But in all fairness, the professor's overconfidence was natural. Back in the 1960's AI researchers tended to think of vision as rather easy, largely because we do it ourselves with no mental effort at all. A game like chess seemed to require much more thought, and there were already programs that could play chess passably well.

And indeed the goal of vision does seem rather straightforward. As the late David Marr of the Massachusetts Institute of Technology (MIT) recently wrote, "Vision is a process that produces from images of the external world a description that is useful to the viewer and not cluttered with irrelevant information" (*1*).

However, the simplicity is deceptive. It is one thing to record an image with a camera; it is quite another thing to understand what that image represents. In the early 1970's AI researchers began to write vision programs in earnest—and began to realize what a horrendous thing vision really is.

First, a real-world image contains an enormous amount of data, much of it irrelevant and all of it subject to noise and distortion. In practice this means that a vision system has to have huge amounts of memory and processing power. If one begins with a high-resolution image measuring 1000 by 1000 pixels—a "pixel" being a single digitized picture element—even some of the simplest procedures require about 100 million operations. The human retina, which has approximately 100 million rods and cones, plus four other layers of neurons, all operating at roughly 100 hertz, performs at least 10 billion calculations per second before the image even gets to the optic nerve. And then, once the image information reaches the brain, the cerebral cortex has more than a dozen separate vision centers to process it. In fact, from studies on monkey brains it has been estimated that vision in one form or another involves some 60 percent of the cortex.

The upshot is that if seeing seems effortless, it is because we do not have to think about it; the whole massive computation is unconscious. If chess seems hard, it is only because we *do* have to think about it.

Second, one has the ironic fact that with all this information, there is still not enough. An image is just the two-dimensional projection of a three-dimensional world; the reverse transformation, from the 2-D image to the 3-D objects, is highly ambiguous. So far as a 2-D image on the retina is concerned, for example, the family cat might as well be carved into the tip of an infinitely long rod directed straight away from the eye. And yet, because we know that cats are not like that we never perceive the poor beast that way. Clearly, a competent vision system needs to "know" about cats, and dogs, and an enormous variety of other things, just to resolve the ambiguities.

Third, an object may only vaguely resemble others of its generic type. Consider a real cat, a porcelain cat, and a cat made out of twisted pipe cleaners: What is it that allows us to recognize them all as cats? In addition, as lighting conditions or viewing angles change, an object may not even resemble itself; consider a cat as seen from the side, and a cat as seen face on. This fact alone makes the commercial "template-matching" vision systems hopelessly inadequate for anything but the carefully controlled environment of a factory.

Finally, there are a myriad of possible objects in the world, and almost as many generic types. Humans can handle them all, in principle. A powerful vision system should be able to do it too.

Laid out like this, the problem of vision might seem hopeless. But, in fact, the computer vision community is surprisingly optimistic. The next few years promise to bring an enormous increase in computational power, largely due to the development of a new class of processors that do their calculations in parallel instead of in series.

But perhaps more important, there is a sense in the community that the "low-level," or "early" part of the vision problem, the perception of 3-D shape